

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-41260

(P2017-41260A)

(43) 公開日 平成29年2月23日(2017.2.23)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 3/01 (2006.01)	G06F 3/01 510	3C707
B25J 5/00 (2006.01)	B25J 5/00 F	5E555

審査請求 有 請求項の数 21 O L (全 42 頁)

(21) 出願番号	特願2016-186918 (P2016-186918)	(71) 出願人	311003754
(22) 出願日	平成28年9月26日 (2016.9.26)		ソフトバンク・ロボティクス・ヨーロッパ
(62) 分割の表示	特願2013-520054 (P2013-520054)		SOFTBANK ROBOTICS EUROPE
原出願日	平成23年7月11日 (2011.7.11)		フランス国、75015・パリ、リュ・デュ・コロネル・ピエール・アビア・43
(31) 優先権主張番号	1056047	(74) 代理人	100071054
(32) 優先日	平成22年7月23日 (2010.7.23)		弁理士 木村 高久
(33) 優先権主張国	フランス (FR)	(72) 発明者	メゾニエ、ブルーノ
(特許庁注：以下のものは登録商標)			フランス、エフ-75005 パリ、リュ・ジョフロワ サンーティレール、57
1. ブルートゥース		(72) 発明者	モンソー、ジェローム
			フランス、エフ-75013 パリ、リュ・ブッサニョー、6

最終頁に続く

(54) 【発明の名称】 自然な対話インターフェースを備えたヒューマノイドロボット、同ロボットを制御する方法、および対応プログラム

(57)【要約】 (修正有)

【課題】 音声および可視認識処理の能力に制限されない、対話者との自然な対話のためのインターフェースを備えたヒューマノイドロボットを提供する。

【解決手段】 ロボットは、音声信号や画像信号を受信する複数のタイプの受信チャネル521、522、523を備える。制御モジュール510は、受信チャネル521、522、523からメッセージを受信し、受信したメッセージの理解の信頼水準を評価する。制御モジュール530は、信頼水準の評価結果に依存するメッセージを生成し、出力チャネル531、532、533出力する。

【効果】 ロボットの対話者との自然な対話を行い、対話の品質を改良する。

【選択図】 図5

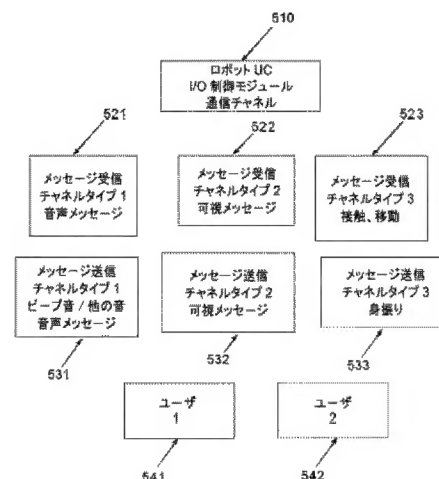


図 5

【特許請求の範囲】

【請求項1】

異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャンネルであってそれぞれが受信チャンネルと送信チャンネルを含むチャンネルの群から選択される2つのチャンネルと、前記チャンネルの入力／出力のための制御モジュールと、を含むヒューマノイドロボットであって、

前記制御モジュールは、第1のチャンネルと第2のチャンネル上で受信／送信されたメッセージを組み合わせる機能とチャンネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される、ヒューマノイドロボットであり、

10

第1のチャンネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第1のタイプのメッセージを受信する受信チャンネルであり、第2のチャンネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される前記第1のタイプとは異なる第2のタイプのメッセージを送信する送信チャンネルであり、

前記制御モジュールは、前記第1のタイプのメッセージを受信する受信チャンネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第2のタイプのメッセージを送信する送信チャンネル上にその内容が前記信頼水準に依存する少なくとも1つの第2のメッセージを生成することができる、ヒューマノイドロボット。

20

【請求項2】

前記チャンネルは、可聴、可視、触知可能、身振りの、位置的、または符号メッセージを送信および／または受信するための通信チャンネル群から選択される、請求項1に記載のヒューマノイドロボット。

【請求項3】

第1の通信チャンネルは音声送信チャンネルであり、第2の通信チャンネルは前記少なくとも1人の対話者による前記ロボットの少なくとも一部の身振りおよび／または位置を受信するためのチャンネルであり、

前記身振りおよび／または位置は前記対話者により前記ロボットに伝えられる入力を表し、

30

前記入力の仕様は、前記第1のチャンネル上で送信された前記メッセージにより、前記ロボットにより前記対話者に対し規定される、請求項2に記載のヒューマノイドロボット。

【請求項4】

前記対話者が前記第2のチャンネル内への前記入力を検証する第3の触知能通信チャンネルをさらに含む、請求項3に記載のヒューマノイドロボット。

【請求項5】

前記第1のチャンネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信されたメッセージを音声認識するためのフィルタを含み、

前記第2のメッセージの前記内容は、前記第1のチャンネル上の前記第1のメッセージの繰り返しの要求と、前記フィルタの前記表現のサブセットの前記第1のチャンネル上で前記対話者により送信される第3のメッセージによる確認の要求と、少なくとも1つの第3のチャンネル上の少なくとも1つの他のメッセージの前記対話者による送信の要求と、からなる要求の群から発見的方法により選択される、請求項2に記載のヒューマノイドロボット。

40

【請求項6】

前記第1と第2のチャンネル上の前記メッセージの半二重モードでの順番を保証するように、前記第1のチャンネル上での聴取開始信号を前記第2のチャンネル上で送信するようにさらに構成される、請求項5に記載のヒューマノイドロボット。

【請求項7】

前記発見的方法は前記期待認識率に基づき決定される閾値に関する実認識率の位置の関

50

数である、請求項5に記載のヒューマノイドロボット。

【請求項8】

前記第3のチャンネルは触覚的受信のためのまたは前記ロボットの一部から身振りを受信するためのチャンネルである、請求項5に記載のヒューマノイドロボット。

【請求項9】

電子メールシステムとインターフェースするためのモジュールであって、前記第1と第2のチャンネルを使用することにより、前記電子メールシステム上のアカウントのホルダが前記ロボットを、前記第2のチャンネル上の電子メールを受信し／読み込み、前記第1のチャンネル上で書き込み／送信し、対話により前記アカウントを管理するためのエージェントとして使用できるようにする、モジュール、をさらに含む請求項2に記載のヒューマノイドロボット。

10

【請求項10】

前記第3のチャンネルは前記第1のチャンネルの前記フィルタの表現の前記リストに対応する対象物の像の可視受信のためのチャンネルであり、

前記像は、前記表現と共に以前に記録された前記対象物の像のデータベースであって前記通信チャンネルの入力／出力のための前記制御モジュールによりアクセス可能なデータベースと比較される、請求項5に記載のヒューマノイドロボット。

【請求項11】

第1の通信チャンネルは可視メッセージを受信するためのチャンネルであり、第2の通信チャンネルは音声メッセージを送信するためのチャンネルであり、

20

前記制御モジュールは、前記第1のチャンネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第2のチャンネル上にその内容が前記信頼水準に依存する少なくとも1つの第2のメッセージを生成することができる、請求項2に記載のヒューマノイドロボット。

【請求項12】

前記第1のチャンネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信された前記メッセージの像を認識するためのフィルタを含み、

前記第2のメッセージの前記内容は、前記第1のチャンネル上の前記第1のメッセージの繰り返しの要求と、前記フィルタの前記表現のサブセットの音声メッセージを受信するための第3のチャンネル上で前記対話者により送信される第3のメッセージによる確認の要求と、少なくとも1つの第4のチャンネル上の少なくとも1つの他のメッセージの前記対話者による送信の要求と、からなる要求の群から発見的方法により選択される、請求項11に記載のヒューマノイドロボット。

30

【請求項13】

前記チャンネルの少なくとも1つは、前記入力と出力のための前記制御モジュールによりマージされた2つのチャンネルの出力を入力として受信するハイブリッドチャンネルである、請求項1に記載のヒューマノイドロボット。

【請求項14】

少なくとも1人の対話者とヒューマノイドロボットとの通信の制御方法であって、

異なる様式を使用する通信チャンネルによるメッセージ送信の少なくとも2つの工程であってそれぞれが受信チャンネルと送信チャンネルを含むチャンネルの群から選択される2つの工程と、

40

前記チャンネルの入力／出力を制御する工程と、を含み、

前記制御工程は、第1のチャンネルと第2のチャンネル上で受信／送信されたメッセージを組み合わせる機能とチャンネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成され、第1のチャンネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第1のタイプのメッセージを受信する受信チャンネルであり、第2のチャンネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセー

50

ジの各タイプのメッセージから選択される前記第1のタイプとは異なる第2のタイプのメッセージを送信する送信チャネルであり、

前記制御工程は、前記第1のタイプのメッセージを受信する受信チャネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第2のタイプのメッセージを送信する送信チャネル上にその内容が前記信頼水準に依存する少なくとも1つの第2のメッセージを生成することができる、方法。

【請求項15】

コンピュータプログラムがコンピュータ上で実行されると請求項14に記載の方法の実行を可能にするプログラムコード命令を含むコンピュータプログラムであって、

前記プログラムは、異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される少なくとも2つのチャネルと、前記チャネルの入力／出力の制御のためのサブルーチンと、を含むヒューマノイドロボットを可能にするように構成され、

前記制御サブルーチンは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される、コンピュータプログラム。

【請求項16】

異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される少なくとも2つのチャネルと、前記チャネルの入力／出力のための制御モジュールと、を含む少なくとも1つのヒューマノイドロボットと少なくとも1人の対話者との間の通信インターフェースを編集および制御する方法であって、

前記制御モジュールは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成され、

前記選択された機能をプログラムする工程をさらに含み、

第1のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される第1のタイプのメッセージを受信する受信チャネルであり、第2のチャネルは可聴、可視、触知可能、身振りの、位置的、または符号メッセージの各タイプのメッセージから選択される前記第1のタイプとは異なる第2のタイプのメッセージを送信する送信チャネルであり、

前記制御モジュールは、前記第1のタイプのメッセージを受信する受信チャネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第2のタイプのメッセージを送信する送信チャネル上にその内容が前記信頼水準に依存する少なくとも1つの第2のメッセージを生成することができる、方法。

【請求項17】

前記選択された機能をプログラムする前記工程は、音声送信チャネルを装った第1の通信チャネルと、前記少なくとも1人の対話者により前記ロボットの四肢上に加えられた少なくとも1つの身振りを受信するためのチャネルを装った第2の通信チャネルと、を規定する少なくとも1つの下位工程と、

前記対話者により前記ロボットに伝えられた前記少なくとも1つの身振りと入力との対応を規定する下位工程と、

前記第1のチャネル上で前記ロボットにより前記対話者に送信される少なくとも1つのメッセージを生成することにより前記入力の仕様を規定する下位工程と、を含む、請求項16に記載の編集および制御する方法。

10

20

30

40

50

【請求項18】

第3の触知通信チャネルを定義し、これにより前記対話者が前記第2のチャネルに対し行われた前記入力を検証する下位工程、をさらに含む請求項17に記載の編集し制御する方法。

【請求項19】

前記選択された機能をプログラムする前記工程は、

音声メッセージを受信するチャネルを装った第1の通信チャネルと音声メッセージを送信するためのチャネルを装った第2の通信チャネルとを規定する少なくとも1つの下位工程と、

前記第1のチャネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価するための機能を規定する下位工程と、

その内容が前記信頼水準に依存する前記第2のチャネル上の少なくとも1つの第2のメッセージの生成を規定する下位工程と、を含む、請求項16に記載の編集および制御する方法。

【請求項20】

コンピュータプログラムがコンピュータ上で実行されると請求項16に記載の方法を実行するためのプログラムコード命令を含むコンピュータプログラムであって、

前記プログラムは、異なる様式に応じた、少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つのチャネルと、前記チャネルの入力／出力を制御するためのサブルーチンと、を含むヒューマノイドロボットをプログラムできるように構成され、

前記コンピュータプログラムは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択され前記ロボットにより実行される少なくとも1つの機能を制御するための前記サブルーチン内でプログラムするためのモジュールを含む、コンピュータプログラム。

【請求項21】

少なくとも1つのパラメータを制御ボックスに渡すことをプログラムするためのモジュールをさらに含む請求項20に記載のコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はヒューマノイドロボットの分野に属する。より正確には、本発明は、ユーザの命令に基づく行為の前記ロボットによる実行と、前記ロボットによる適切な回答の提供と、より一般的なやり方では前記ロボットとその対話者または複数の対話者との「ヒューマノイド関係」の確立と、を可能にするように、この種のロボットとの対話のためのインターフェースをプログラムし採用する方法に適用される。

【背景技術】

【0002】

ロボットは、人間の外観および機能：頭、胴、両腕、任意選択的に両手、両脚、両足などのいくつかの属性を有する限りヒューマノイドと称されることがある。外観以外では、ヒューマノイドロボットが達成できる機能は、運動を行い、話し、「推論する」ためのその能力に依存する。ヒューマノイドロボットは四肢または頭を使って歩行し、身振りをすることができる。ヒューマノイドロボットが行える身振りの複雑さは引き続き増加している。

【0003】

いくつかのロボットは環境からの刺激に応じて話すことができる。相互作用の可能性を著しく豊かにするいくつかのロボットと人間との対話のための機能を開発できるようにする音声認識および合成ツールの開発もなされてきた。このような言葉を使用する人間－ロ

10

20

30

40

50

ボットインタフェースは、特に米国特許出願公開第2009/287678号明細書だけでなく米国特許第7,711,569号明細書により開示されている。

【0004】

これらの従来技術文書では、音声認識に固有の不完全性は、データベースへのアクセスを必要とする意味論的および／または文脈論的支援に訴えることにより、学習により、認識の疑い（低認識信頼区間、偽陽性、偽陰性など）を解消できるように大きな計算資源を利用することにより、緩和される。これらの手段の利用は、移動運動などのその最重要処理を管理するようにその計算資源が経済的でなければならない多機能ヒューマノイドロボットの場合は、適切でない。

【0005】

前記ロボット上に埋め込まれた計算資源を使用することにより簡単かつ効果的やり方で、不完全なままとなるセンサとソフトウェアにより行われる認識に関する疑いを解消できるヒューマノイドロボットを有することができれば有利だろう。

【発明の概要】

【課題を解決するための手段】

【0006】

本発明は、回答の確認の自然なモードを使用するヒューマノイドロボットとの対話のためのインターフェースを提供することによりこの問題を解決する。

【0007】

この目的のため、本発明は、異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つのチャネルと、2つのチャネルの入力／出力のための制御モジュールと、を含むヒューマノイドロボットを開示する。制御モジュールは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づき前記ロボットにより受信されたメッセージの理解を向上させるように構成される。

【0008】

有利には、前記通信チャネルは、可聴、可視、触知可能メッセージ、ロボットの少なくとも一部分の移動および／または位置に関するメッセージ、およびデジタルメッセージを送信および／または受信する為の通信チャネルの群から選択される。

【0009】

有利には、第1の通信チャネルは音声送信チャネルであり、第2の通信チャネルは少なくとも前記1人の対話者によるロボットの少なくとも一部分からの移動および／または位置を受信するためのチャネルである。前記移動および／または位置は、対話者によりロボットに伝えられる入力を表し、入力の前記仕様は、第1のチャネル上で送信されたメッセージによりロボットにより対話者に対し規定される。

【0010】

有利には、本発明のロボットはさらに、対話者が第2のチャネルに対し行われた入力を検証する第3の触知通信チャネル(tactile communication channel)を含む。

【0011】

有利には、第1の通信チャネルは音声メッセージを受信するためのチャネルであり、第2の通信チャネルは音声メッセージを送信するためのチャネルであり、前記制御モジュールは、前記第1のチャネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価し、前記第2のチャネル上にその内容が前記信頼水準に依存する少なくとも1つの第2のメッセージを生成することができる。

【0012】

有利には、第1のチャネルは、それぞれが期待認識率と関連付けられた表現のリストに

10

20

30

40

50

より、受信されたメッセージを音声認識するためフィルタを含み、第2のメッセージの内容は、第1のチャンネル上の第1のメッセージの繰り返しの要求と、フィルタの表現のサブセットの第1のチャンネル上で対話者により送信される第3のメッセージによる確認の要求と、少なくとも1つの第3のチャンネル上の少なくとも別のメッセージにおける対話者による送信の要求と、からなる要求の群から発見的方法により選択される。

【0013】

有利には、本発明のロボットは、第1と第2のチャンネル上のメッセージの半二重モードでの順番を保証するように、第1のチャンネル上での聴取開始信号(s i g n a l o f s t a r t o f l i s t e n i n g)を第2のチャンネル上で送信するように構成されている。

10

【0014】

有利には、前記選択発見的方法は、期待認識率に基づき決定される閾値に関する実認識率(r e a l r e c o g n i t i o n r a t e)の位置の関数である。

【0015】

有利には、前記第3のチャンネルは、触知的受信のための、またはロボットの一部分から移動を受信するためのチャンネルである。

【0016】

有利には、本発明のロボットはさらに、電子メールシステムとインターフェースするためのモジュールを含む。この前記このモジュールは、前記第1と第2のチャンネルを使用することにより、前記電子メールシステム上のアカウントのホルダが前記ロボットを、第2のチャンネル上で電子メッセージを受信し／読み、第1のチャンネル上で電子メッセージを書き／送り、対話により前記アカウントを管理するためのエージェントとして使用できるようにする。

20

【0017】

有利には、前記第3のチャンネルは、第1のチャンネルのフィルタの表現のリストに対応する対象物の像の可視受信(v i s u a l r e c e p t i o n)のためのチャンネルであり、前記像は、前記表現と共に以前に記録された前記対象物の像のデータベースであって前記通信チャンネルの入力／出力のための前記制御モジュールによりアクセス可能なデータベースと比較される。

【0018】

30

有利には、第1の通信チャンネルは可視メッセージ(v i s u a l m e s s a g e s)を受信するためのチャンネルであり、第2の通信チャンネルは音声メッセージを送信するためのチャンネルであり、前記支援制御モジュールは、前記第1のチャンネル上で受信される第1のメッセージの前記ロボットによる理解の前記信頼水準を評価し、その内容が信頼水準に依存する前記第2のチャンネル上の少なくとも1つの第2のメッセージを生成することができる。

【0019】

有利には、第1のチャンネルは、それぞれが期待認識率と関連付けられた表現のリストにより、受信されたメッセージの像を認識するためのフィルタを含み、前記第2のメッセージの内容は、第1のチャンネル上の前記第1のメッセージの繰り返しの要求、フィルタの表現のサブセットの音声メッセージを受信するための第3のチャンネル上で対話者により送信される第3のメッセージによる確認の要求、少なくとも1つの第4のチャンネル上の少なくとも1つの他のメッセージの対話者による送信の要求を含む、要求の群から発見的方法により選択される。

40

【0020】

有利には、チャンネルの少なくとも1つは、入力と出力のための前記制御モジュールによりマージされた2つのチャンネルの出力を入力として受信するハイブリッドチャンネルである。

【0021】

本発明はまた、少なくとも1人の対話者とヒューマノイドロボットとの通信の制御方法

50

を開示する。本方法は、異なる様式を使用する通信チャネルによるメッセージ送信の少なくとも2つの工程であってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つの工程と、チャネルの入力／出力を制御する工程と、を含む。制御工程は、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。

【0022】

本発明はまた、コンピュータプログラムがコンピュータ上で実行されると本発明の方法の実行を可能にするプログラムコード命令を含むコンピュータプログラムを開示する。本プログラムは、異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つのチャネルと、チャネルの入力／出力の制御のためのサブルーチンと、を含むヒューマノイドロボットを可能にするように構成される。制御サブルーチンは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。

10

【0023】

本発明はまた、異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つのチャネルと、2つのチャネルの入力／出力のための制御モジュールと、を含む少なくとも1つのヒューマノイドロボットと少なくとも1人の対話者との通信インターフェースを編集し制御する方法を開示する。制御モジュールは、第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能とチャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能とを含む機能の群から選択された少なくとも1つの機能の実行に基づきロボットにより受信されたメッセージの理解を向上させるように構成される。本方法は、選択された機能をプログラムする工程をさらに含む。

20

30

【0024】

有利には、前記選択された機能をプログラムする前記工程は、音声送信チャネルを装った第1の通信チャネルと、前記少なくとも1人の対話者によりロボットの少なくとも一肢の移動を受信するためのチャネルを装った第2の通信チャネルと、を規定する少なくとも1つの下位工程と、対話者によりロボットに伝達される前記移動と入力との対応を規定する下位工程と、第1のチャネル上でロボットにより対話者に送信される少なくとも1つのメッセージを生成することにより前記入力の仕様を規定する下位工程と、を含む。

【0025】

有利には、本発明の編集及び制御する方法はさらに、第2のチャネルに対し行われた入力を対話者が検証する第3の触知通信チャネルを規定する下位工程を含む。

40

【0026】

有利には、前記選択された機能をプログラムする前記工程は、音声メッセージを受信するためのチャネルを装った第1の通信チャネルと音声メッセージを送信するためのチャネルを装った第2の通信チャネルとを規定する少なくとも1つの下位工程と、前記第1のチャネル上で受信された第1のメッセージの前記ロボットによる理解の信頼水準を評価するための機能を規定する下位工程と、その内容が前記信頼水準に依存する前記第2のチャネル上の少なくとも1つの第2のメッセージの生成を規定する下位工程と、を含む。

【0027】

本発明はまた、コンピュータプログラムがコンピュータ上で実行されると本発明の方法を実行するためのプログラムコード命令を含むコンピュータプログラムを開示する。コン

50

コンピュータプログラムは、異なる様式に応じた少なくとも1人の対話者とのメッセージの自然なコミュニケーションのための少なくとも2つのチャネルであってそれぞれが受信チャネルと送信チャネルを含むチャネルの群から選択される2つのチャネルと、チャネルの入力／出力を制御するためのサブルーチンと、を含むヒューマノイドロボットをユーザがプログラムできるように構成される。本コンピュータプログラムは、ロボットにより実行される少なくとも1つの機能であって第1のチャネルと第2のチャネル上で受信／送信されたメッセージを組み合わせる機能と、チャネル上で受信された第1のメッセージに基づき生成された第2のメッセージを送信する機能と、を含む機能の群から選択される少なくとも1つの機能を制御するためのサブルーチン内でプログラムするためのモジュールを含む。

10

【0028】

有利には、本発明のコンピュータプログラムはさらに、少なくとも1つのパラメータを制御ボックスに渡すことをプログラムするためのモジュールを含む。

【0029】

本発明のインターフェースはさらに、例えば音声認識が何らかの有効性を持つことができるには周囲騒音が高過ぎる場合に対話が実行される環境に容易に適合化され得る多様式(multimodal)確認モードの提供という利点を提示する。したがってユーザには、接触、身振り、または特定の色または形状のデジタル符号の表示により曖昧な回答を置換／確認するようことを勧めることができる。したがってユーザは、ユーザがコンピュータに向かい合った場合またはインテリジェント電話またはタッチパッドを使用する場合に使用し慣れている従来のインターフェースを直観的やり方で置換またはエミュレートできるようにする手段を思うままに有する。

20

【0030】

さらに、ロボットの表現のモード自体は、対話者の注意を維持するようにおよび提供される回答に関する感情またはヒントを対話者に伝えるように、特にはイントネーション、凝視、身振りを組み合わせることにより、多様式となることができる。さらに、人間同士の自然なコミュニケーションのモードに近づくことにより、本発明のインターフェースは、認識システムの結果を改良すること、「真の仮想性」に埋没されるユーザの経験(すなわち、物理的人間の化身との対話)の質を向上させること、に貢献する。

【0031】

本発明はまた、その設計者により想定されなかったロボットの用途に特に適合化された新しい相互作用シナリオを極めて容易におよびほんの短い時間で作成できるようにするインターフェースを開発するための人間工学的および多目的環境を提供する。

30

【0032】

本発明は、いくつかの例示的实施形態の以下の説明とその添付図面からより良く理解され、その様々な特性と利点が浮かび上がることになる。

【図面の簡単な説明】

【0033】

【図1】本発明のいくつかの実施形態におけるヒューマノイドロボットの物理的アーキテクチャの図である。

40

【図2】本発明の実施形態のいくつかにおける実施に有用なセンサを含むヒューマノイドロボットの頭を示す。

【図3】本発明のいくつかの実施形態におけるロボットの機能の制御を可能にするハイレベルソフトウェアのアーキテクチャの図である。

【図4】本発明のいくつかの実施形態におけるロボットの行動／相互作用を編集しプログラムするための機能アーキテクチャの図である。

【図5】本発明のいくつかの実施形態においてヒューマノイドロボットが受ける応答／刺激のヒューマノイドロボットにより与えられる解釈を改良するために一般的なやり方で適用される処理動作の機能フローチャートである。

【図6】本発明のいくつかの実施形態におけるロボットの行動／相互作用のプログラミン

50

グ論理チャートである。

【図7 a - c】本発明のいくつかの実施形態における多様式インターフェース (m u l t i m o d a l i n t e r f a c e) の相互作用の論理的および時間的組み合わせを示すタイムチャートを表す。

【図8 a - e】本発明の一実施形態における対話言語を変更するための2進選択および選択肢によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。

【図9 a - e】本発明の一実施形態における対話言語を変更するための選択肢のリストとからの選択によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。

10

【図10 a - d】本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。

【図11 a - b】本発明の一実施形態における選択のリストの選択肢を置換または補足することができるようにし、いくつかの選択肢同士を比較する新しい音声認識試験を実行できるようにする一連の画面を表す。

【図12 a - d】本発明の一実施形態における質問のものとは異なる言語で選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。

【図13 a - d】本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験の閾値を検証／修正できるようにする一連の画面を表す。

20

【発明を実施するための形態】

【0034】

図1に、本発明の一実施形態におけるヒューマノイドロボットの物理的アーキテクチャを示す。このようなロボットは、特に2009年10月15日公開の国際公開第2009/124951号パンフレットに開示されている。このプラットフォームは、本発明に至る改良のためのベースとして役立った。後の説明では、このヒューマノイドロボットはこの総称用語またはその商標N A O (商標) で呼ばれることがあるが、これにより参照の一般性が修正されることはない。

【0035】

このロボットは、センサと関節を駆動するアクチュエータとを制御するためのタイプ110の電子カードを約2ダース含む。図に示すカード110は左足を制御するものである。本アーキテクチャの長所の1つは、関節を制御するカードはその大部分が交換可能であるということである。関節は通常、少なくとも2つの自由度したがって2つのモータを有する。各モータは角度で駆動される。関節はまた、いくつかの位置センサ、特にM R E (磁気回転符号器: M a g n e t i c R o t a r y E n c o d e r s) を含む。電子制御カードは市販のマイクロコントローラを含む。これは、例えばM i c r o c h i p 社のD S P I C (商標) であってよい。これはD S P に結合された16ビットMCUである。このMCUは、1ミリ秒のループスレービングサイクル (l o o p e d s l a v i n g c y c l e) を有する。ロボットはまた、他のタイプのアクチュエータ、特にその色と強度がロボットの感情を伝えることができるL E D (発光ダイオード) を含むことができる。後者はまた、他のタイプの位置センサ、特に慣性ユニット、F S R (接地圧センサ) 等を含むことができる。

30

40

【0036】

頭160は、ロボットの知能、特に、ロボットが割り当てられた使命、特に本発明の枠組み内でゲームへの参加を達成できるようにするハイレベル機能を実行するカード130を含む。但し、カード130は、ロボット内の他のどこか例えば胴内に置かれてもよいであろう。しかしながらこの場所は、頭が着脱可能な場合、これらのハイレベル機能を交換する、したがって特にロボットの知能としたがってその使命を極めて速やかに完全に変更できるようにすることが分かるであろう。または逆に、同じ人工知能を保持する一方で1つのボディを別のボディと交換する (例えば、欠陥のあるボディを欠陥の無いボディと)

50

ことができるようにする。頭はまた、特に言葉または画像を処理するための、または、またWAN (Wide Area Network) 広域ネットワーク上の遠隔通信を確立するためにポートを開くために必要なコード化などのサービス入力／出力を処理するための特殊なカードを含むことができる。カード130のプロセッサは市販のx86プロセッサであってよい。AMD社のGeode (商標) (32ビット、500MHz)などの低消費電力プロセッサが好適なやり方で選択されることになる。カードはまた、一組のRAMとフラッシュメモリを含む。このカードはまた、通常はWi-Fi、Wi-Maxトランスミッションレイヤ上で、随意的にはVPN内の随意的にカプセル化された標準プロトコルを有するデータの移動体通信用公衆ネットワーク上で、ロボットの外部(行動サーバー(behaviors server)、他のロボットなど)との通信を管理する。プロセッサは通常、標準OSにより駆動され、これにより通常のハイレベル言語(C、C++、Python等)、またはハイレベル機能をプログラムするためのURBI(ロボット工学のための特殊なプログラミング言語)などの人工知能のための特定言語を使用できるようにする。カード120はロボットの胴内に收容される。これは、カード130により計算された命令のカード110への送信を保証する計算機が位置する場所である。このカードをロボット内の他のどこかに收容することも可能であろう。しかし胴内の場所は、頭の近くでかつ四肢の十字路に位置し、したがってこれによりこのカード130をカード120とカード110にリンクする接続配置を最小化できるので有利である。このカード120の計算機もまた市販のプロセッサである。これは100MHzクロックのARM9(商標)タイプの32ビットプロセッサであると有利である。プロセッサの型、オン／オフボタンに近いその中央位置、電源の制御へのリンクは、ロボットの電源(スタンバイモード、非常停止等)を管理するように好適に適合化されたツールとなる。カードはまた、一組のRAMとフラッシュメモリを含む。

10

20

【0037】

3つのレベルを有するこのアーキテクチャは、ロボットが、連携動作と、センサの読み取りと同時にその環境内で発せられた言葉または符号を解釈し、それに反応または応答するような他の行為と、を実行できなければならない本発明の実施形態には特に有利である。

【0038】

図2aと図2bに、本発明の実施形態のいくつかにおける発明の実施に有用なセンサを含むヒューマノイドロボットの頭の末端図とプロフィール図をそれぞれ表す。

30

【0039】

図1の頭160は、本発明の実施に有用な感知能力と表現能力をロボットに備えさせるように頭200a、200bとして改善されている。

【0040】

NAOは、例えばKingstate Electronics Corpにより提供されるKEEG1540PBL-Aと呼ばれる4つの全指向性マイクロホン211a、212a、213a、214a、すなわち前部には211a、裏側には214a、頭の各側面に212a、213aを備える(図2bも参照されたい)。これらは頭の内部に分散されるので、外部にアクセスするための孔だけが同図では視認可能である。マイクロホンにより行われる音声捕捉に基づき、音声分析および認識システム(例えば、Acapela(商標)社により提供されるシステムBabEAR(商標))は、本明細書内の後で提示される適切なインターフェースを有するユーザが自身の用語により豊かにすることができる所定の単語のコーパスを認識する。これらの単語は、ユーザの好みの行動、特にロボットにより解釈される質問に対する回答をトリガできるようにする。ソフトウェア環境は本明細書の後で示されるようにいくつかの言語を支援する。NAOはまた、音声の源を検出することができこれにより何人かの対話者間の曖昧さを解消できるようにする。

40

【0041】

NAOは、1秒当たり最大30個の画像を捕捉することができる2台のCMOS 640×480カメラ(220a)、例えばOV760と呼ばれるOmni-vision(商

50

標)ブランドのカメラ(CMOS 1/6インチセンサ:3.6μmの画素)を通して見る。額のレベルに置かれた第1のカメラはその水平線に向けられ、口のレベルに置かれた第2のカメラはその直接環境を調査する。ソフトウェアは、NAOが見たものの写真および映像ストリームをも読み出すことができるようにする。その環境を感知し解釈するために、NAOは、ボールとより複雑な対象物とを捜し出すために、対話者を認識できるようにする一組の顔および形状の検出と認識アルゴリズムを搭載する。

【0042】

NAOは、その頭蓋の頂部上に置かれた、特にこのアプリケーションのために出願人により例えば3つの区画に分割されて開発された容量センサ(230a)を備える。特定のアプリケーションでは、4区画以上を設けることができるであろう。したがって、接触により(例えば、アプリケーションにより規定された行動のトリガを可能にする一連のボタンを押すことにより)、本発明の枠組み内で各ボタンに関連する異なる回答であってよい情報をNAOに与え、提案された選択のリストをスクロールして進み、ヘルプメニュー等にアクセスすることが可能である。システムには、接触があるかどうかを示すLEDが付随する。

10

【0043】

NAOは、例えば明細書の後で説明されるまたはWebサイトまたはRSSストリームから検索されるモードに応じてプログラムされるその格納スペース内に局所的に存在する任意のテキストファイルを大声で読むことにより自身を表現することができる。頭の各側面上に配置された2つの拡声器210bを備えることにより、その音声合成システム(例えば、AcapelaのAcapela Mobility)はパラメータ化可能であり、これにより特に音声の速度および/またはトーンの修正を可能にする。

20

【0044】

NAOに音楽ファイルを送りそれを演奏させることが可能である。NAOは例えば、wavおよび、mp3フォーマットを受け入れ、これにより本発明の枠組み内で、音声回答の付属としてまたはその代替として音楽回答または専用音声を提供できるようにする。音楽ファイル以外のフォーマットを受け入れてもよい。

【0045】

図3は、本発明の一実施形態におけるロボットの機能の制御を可能にするハイレベルソフトウェアのアーキテクチャの図である。

30

【0046】

このタイプのソフトウェアアーキテクチャは、特に2009年10月15日公開の国際公開第2009/124955号パンフレットに開示されている。このアーキテクチャは、ロボットとPCまたは遠隔場所との間の通信を管理するための、そして本発明の実施のために必要なソフトウェアインフラストラクチャを提供するソフトウェアを交換するための、基本機能を含む。このアーキテクチャは、特定のアプリケーションにおいて使用されるソフトウェア機能の特定の言及無しに包括的なやり方で以下に説明される。これらの機能は本発明のロボットの行動を管理するための任意の他のソフトウェア機能として処理されることが分かる。

【0047】

40

図3では、例えば移動性理由のために無線リンクにより第1の遠隔端末装置TD1と通信する第1のヒューマノイドロボットRH1が極めて図式的に表される。遠隔端末装置という表現は、通信ネットワークを手段として、このタイプのヒューマノイドロボットRH1に専用化されたウェブサービスSWへのアクセスを提供するサーバプラットフォームPFSから離れた端末を意味するように意図されている。

【0048】

当然、システムの要素間の通信リンクは有線ベースであってもよく、移動端末は変形形態として携帯電話または携帯型コンピュータであってもよい。

【0049】

第2のヒューマノイドロボットRH2は、例えばヒューマノイドロボットRH2の移動

50

性を妨げないように無線リンクによっても第2の遠隔端末装置TD2と通信する。

【0050】

遠隔端末装置TD1、TD2とサーバプラットフォームPFSは、通信ネットワークRCを手段としてネットワーク内でリンクされる。遠隔端末装置TD1、TD2だけでなくサーバプラットフォームPFSのウェブサービスのために、またヒューマノイドロボットRH1、RH2のために、少なくとも1系列の命令を含む少なくとも1つのモジュールに専用化された単一のそれぞれの連結モジュールB5、B2、B4、B1、B3はプロセッサにより実行されるソフトウェア機能を実施する。連結モジュールB5、B2、B4、B1、B3のそれぞれのモジュールM51、M52、M21、M22、M41、M42、M11、M12、M31、M32は、この例では、連結モジュール当たりの数は2で表されるがこの数は連結モジュール毎に異なり任意であってよい。

10

【0051】

次に、第1のヒューマノイドロボットRH1を所有する第1の遠隔端末装置TD1のユーザにより想定されるシステムの動作の全く非限定的な実施例を説明する。ユーザは例えば、第1の遠隔端末装置TD1上に搭載されたまたは第1の遠隔端末装置TD1からサーバプラットフォームPFS上にアクセス可能なソフトウェアアプリケーションによりある数の機能をロボットを介し実行することができる。

【0052】

例えば、ユーザは、ソフトウェアアプリケーションのグラフィックツールにより、ロボットが10秒間歩いてから「皆さん、今日は」と言うロボットのアプリケーションを簡単に実行する。このアプリケーションは、例えばモジュール(例えば、モジュールM11)の形式で第1のヒューマノイドロボットRH1内にダウンロードされ、次に第1の遠隔端末装置TD1を手段としてユーザによりトリガされる。

20

【0053】

第1のヒューマノイドロボットRH1は、「歩行」機能を最初に使用しなければならないモジュールM11をトリガする。次に、モジュールM11は、モジュールM11がリンクされる連結モジュールB1に要求する接続インターフェース/機能呼び出しモジュールまたはプロキシP1を使用する。連結モジュールB1は、モジュール内に有する呼び出される機能の場所でネットワーク連結モジュールが要求に応答するまで、それ自身のモジュールに向けられた要求と、反復的やり方でこの動作を繰り返すそれが直接リンクされるネットワークと結合するためのモジュール(子連結モジュール)に向けられた要求と、を行う。この要求に対する応答はまた、プロキシP1に直接リンクされた連結モジュールB1がこの機能を接続し呼び出す必要があるまで親連結モジュールにより反復的やり方で送信される(逆方向に)。例えば、要求された歩行機能は第2の遠隔端末装置TD2のモジュールM41内に置かれる。次に、連結モジュールB4は、例えばロボットが歩行する継続時間を表す秒単位の整数型の継続時間パラメータと、ロボットの排他的または非排他的歩行(すなわちロボットは歩行しながら別の行為を行うことが許容されるか否か)を表すブーリアン型の排他的パラメータと、を含む「歩行」機能の呼び出しのパラメータを返した。この例では10秒歩いた後に話すことが望ましいので、この例では歩行機能は、10秒に等しい継続時間パラメータと1秒に等しい排他的パラメータで呼び出される。

30

40

【0054】

したがって接続インターフェース/呼び出しモジュールP1は、所望のパラメータを有する「歩行」機能への接続とその呼び出しとを、あたかもこの機能がローカルに存在するかのよう遠隔的に行うことができる。接続インターフェースと機能呼び出しモジュールは、異なる端末またはサーバ上に配置されたモジュールの機能を呼び出すことができる相互通信ソフトウェアを使用する。この機能は、呼び出しモジュールのものとは異なるコンピュータ言語の一連の命令により書くことができる。プロキシは、例えばSOAP相互通信ソフトウェアを使用する。これによりプラットフォーム間および言語間通信アーキテクチャを生成する。

【0055】

50

この非局在化(d e l o c a l i z e d)「歩行」機能が行われると、モジュールM11は「話す」機能と呼び出さなければならない。別の接続インターフェースと機能呼び出しモジュールまたはプロキシP2は、モジュールM11がリンクされる連結モジュールB1に要求をする。連結モジュールB1は、当初、格納された一連の命令の形式で実行される機能を手段として、それ自身のモジュールM11とM12に向けられた要求を行い、モジュールM11とM12は例えば、モジュールM12内のこの「話す」機能の有無を返すことになる。連結モジュールB1は接続インターフェース／機能呼び出しモジュールP2に通知し、次にこのモジュールPは、ローカル呼び出し型の呼び出しにより、パラメータとして例えば「今日は」と読まれるテキストを有するモジュールM12の「話す」機能を直接呼び出すことができる。このパラメータは連結モジュールB1によりプロキシP2に送信される。さらに、システムは、ヒューマノイドロボットRH1のこの事例では、外部事象を受信すると前記パラメータの値を更新し前記格納されたパラメータの1つの更新を先の要求に応じてモジュールに通知するように適合化された移動端末の状態を表すパラメータの格納／管理モジュールSTM(「短期記憶: S h o r t T e r m M e m o r y」の呼称)を含む。したがって予め警告を受けたモジュールは、通知されたパラメータの変更の関数として行為を引き受けることができることになる。

10

【0056】

先に説明した例に関連して、例えば、格納／管理モジュールSTMはロボットRH1の動き検出器により検出された誰かの外観を表すパラメータの状態を格納することができる。このパラメータがロボットの直接環境内に誰もいないことを表す状態からロボットの直接環境内の誰かの存在を表す状態に移ると、モジュールM11により先に行われた要求に応じて、格納／管理モジュールSTMは事象または信号を介しこの値の変化を予め警告する。次にモジュールM11は、例えば、先に説明した逐次的トリガ(「歩行」と「話す」機能)を自動的にトリガすることができる。

20

【0057】

図3の例では、格納／管理モジュールSTMは、遠隔端末装置TD1の一部を形成するが、変形形態として、他の遠隔端末装置TD2、サーバープラットフォームPFS、またはヒューマノイドロボットRH1またはRH2の一部を形成することができる。

【0058】

格納／管理モジュールのSTMはまた、メモリ内にそれぞれの基準時間間隔にわたるいくつかのパラメータの時間的進展を格納することができる。したがってシステムのモジュールはさらに、ある継続時間からのパラメータの値の進展にアクセスし、これら取るべき行為の進展を考慮することができる。

30

【0059】

変形形態として、呼び出される機能のモジュールは、ヒューマノイドロボットRH1、RH2上の、または通信ネットワークRCの遠隔端末装置TD1、TD2上のサーバープラットフォームPGS上に置かれてもよい。

【0060】

したがって本発明は、機能に対しローカルまたは遠隔呼び出しを行うかどうかに関わらず、ネットワーク全体にプログラムを分散させ移動端末の同一動作を可能にする。

40

【0061】

さらに、本アーキテクチャはまた、移動端末の状態を表す一組の格納パラメータを有するようにでき、いくつかの行為を自動的にトリガするようにこの状態の進展を考慮できるようにする。

【0062】

さらに、格納／管理モジュールはまた、所定時間間隔中にパラメータの値の進展を記録することができ、これによりモジュールがこれらのパラメータの進展のログにアクセスできるようにする。

【0063】

ロボットのインターフェースを管理するためのNAOQIと名付けられたオペレーティ

50

ングシステムを構成するこれらの通信および格納機能は本発明の実施のために特に有用である。

【0064】

図4は、本発明の一実施形態におけるロボットの行動を編集しプログラムするための機能アーキテクチャの図である。このようなアーキテクチャについては2010年5月25日出願の欧州特許出願第2010/057111号明細書により説明されている。前記アーキテクチャを実装できるようにするヒューマノイドロボットの行動を編集しプログラムするためのソフトウェアは、Choregraphe（商標）と商業的に命名されており、上記特許文献の一般性を損なうことなくその総称または商標のいずれかで呼ぶことができる。

10

【0065】

このアーキテクチャにより制御されるロボットは頭、胴、四肢を有するヒューマノイドロボットであってよく、その各部品は関節で繋がれ、各関節は1つまたは複数のモータにより制御される。このアーキテクチャは、システムのユーザが、仮想ロボット上でシミュレートされ有線ベースまたは無線のリンクによりシステムにリンクされた実ロボット上で実行される行動を生成することによりこのようなロボットを制御できるようにする。

【0066】

これは、そうするようにプログラムされたコンピュータの画面上の行動（例えば、歩く－真直ぐに、右へ、または左へn歩；「今日は」－頭上への片腕の動き；言葉等）と（所与の角度の頭、一肢の）動きとを視認すること、シミュレートすること、実行することを伴う。

20

【0067】

図4は、時間的次元を有する事象によりトリガされる制御の表現を例示する処理工程のフローチャートである。事象によりトリガされる制御は、本発明の意味論では「ボックス」または「制御ボックス」410により表される。ボックスは、この後定義される以下の要素の1つまたは複数を含むことができるツリー状のプログラミング構造である。

- －フレーム420の「タイムライン」または時間軸、
- －「図表」またはフローチャート470、
- －スクリプト490。

【0068】

制御ボックスは通常、本明細書の後で詳述されるように、事象情報項目を1つのボックスから別のボックスに通常は送信する接続部により相互にリンクされる。いかなるボックスも、ロボットの行動／動きのシナリオを初期化する「ルートボックス」またはルートに直接または間接的にリンクされる。

30

【0069】

フレーム420の時間軸は、フレームの前記時間軸が挿入されるボックス内に定義されたロボットの行動と動きとが受ける時間的制約を表す。以下の説明と特許請求範囲では、プログラミングの世界において同じ意味を持って一般的に受け入れられている、タイムラインという用語を使用する。したがってタイムラインはボックスの行動と動きの同期を行う。タイムラインは、毎秒フレーム数（FPS：Frames Per Second）の表現で定義された進捗速度に関連付けられたフレームに細切れにされる。各タイムラインのFPSはユーザによりパラメータ化可能である。デフォルト設定では、FPSは任意の値（例えば、15FPS）で固定されてもよい。

40

【0070】

タイムラインは次のものを含むことができる。

－それぞれが1つまたは複数の行動キーフレーム（Behavior Key Frame）または「主行動フレーム（main behavior Frames）」450を含む1つまたは複数の行動レイヤ（Behavior Layer）430であって、上記行動キーフレーム自体が、行動レイヤまたはタイムラインを通過することなく、実際はより高いレベルのボックスにも直接取り付けられることができるボックスの組である1

50

つまたは複数のチャートまたは「フローチャート」470を含むことができる、行動レイヤ430、

－それぞれが、1つまたは複数の運動画面480を含むことができる1つまたは複数の運動キーフレームまたは「主運動フレーム」460を含む、1つまたは複数の運動レイヤ440。

【0071】

行動レイヤは、一組のロボットの行動または主行動フレームを定義する。いくつかの行動レイヤが同一のボックス内に定義されてもよい。これらは次に、ボックスのタイムラインにより同期されたやり方で進行するようにプログラムされる。

【0072】

行動レイヤは1つまたは複数の主行動フレームを含むことができることになる。主行動フレームは、歩く（「歩行」）、話す（「言う」）、音楽を演奏する（「音楽」）などのロボットの行動を定義する。一定数の行動は、本明細書の後で詳述されるようにライブラリからの単純な「ドラッグアンドドロップ」でユーザにより直接挿入されるように本発明のシステム内に予めプログラムされる。各主行動フレームは、タイムライン内に挿入されるフレームの開始であるトリガ事象により定義される。主行動フレームの終了は、それに続く別の主行動フレームが挿入される限りにおいてだけ、または終了事象が定義される場合に定義される。

【0073】

運動レイヤは、ロボットの関節のモータの動きをグループ化する1つまたは複数の逐次的主運動フレーム（*successive main motion Frames*）によりプログラムされる一組のロボットの運動を定義する。実行されるこれらの動きは、本明細書の後で詳述されるアニメーション画面上の行為によりプログラムされてもよい前記モータの到来角度位置（*angular positions of arrival*）により定義される。同一のボックスのすべての主運動フレームはボックスのタイムラインにより同期される。主運動フレームは到来フレーム（*arrival Frame*）により定義される。開始フレームは、前の主運動フレームの終りのものまたはボックスの開始事象のものである。

【0074】

主行動フレームと主運動フレームは主行為フレーム（*main action Frame*）の一般名で呼ばれる。

【0075】

同じタイムラインに取り付けられているという条件で、いくつかの主行為フレーム（行動、または運動の）を並列に実行することが可能である。

【0076】

フローチャートは後で詳述されるように相互に接続された一組のボックスである。ボックスのそれぞれは次に、新しい行動または運動レイヤが取り付けられる他のタイムラインを含むことができる。

【0077】

スクリプトはロボットにより直接実行可能なプログラムである。本発明の枠組み内では、スクリプトは好ましいやり方でC++言語で書かれる。スクリプトを含むボックスはいかなる他の要素も含まない。

【0078】

ソフトウェアは、Windows（商標）、Mac（商標）、またはLinux（商標）オペレーティングシステムを使用することにより、PC上またはパーソナルコンピュータ型の別のプラットフォーム上に埋め込まれてもよい。

【0079】

本発明のヒューマノイドロボットは通常、Choregraphe（商標）ソフトウェアを使用することにより人間と相互作用できるようにプログラムされることになる。この開発アーキテクチャにより可能になる時間的および行動的論理の組み合わせは、本発明の

10

20

30

40

50

実施のために特に有利である。本明細書の後で説明される一定数のツールは、本発明の枠組み内での自然な対話のためのインターフェースを有するヒューマノイドロボットを実現するために特に開発された。

【 0 0 8 0 】

図5は、本発明のいくつかの実施形態においてヒューマノイドロボットが受ける応答／刺激のヒューマノイドロボットにより与えられる解釈を改良するために一般的なやり方で適用される処理動作の機能フローチャートである。

【 0 0 8 1 】

長期にわたって、人間は機械と相互作用する多様な手段を開発してきた。これらの手段はテクノロジーの進化に従うので、常により効果的となる。すべての場合において、効果的であるためには、相互作用はそれが何であれプラットフォームとユーザの必要性とに適合化されなければならない。

10

【 0 0 8 2 】

したがってグラフィックインターフェースとウィンドウ環境は、例えばテキスト領域（テキストボックス）、OK／キャンセルボタン、チェックマークされるボックス（チェックボックス）、ラジオボタン、または複合ボックス（コンボボックス）などのグラフィックインターフェース部品とも呼ばれる一定数のインターフェース要素（GUI要素すなわちグラフィックユーザインターフェース要素）をユーザの自由意思で配置する。グラフィックインターフェースに適合されたこれらの要素は、従来の画面タイプのいかなる視覚的フィードバックも原理的に提供しないヒューマノイドロボット上などでは使用することができない。今や、ロボットとのやりとりは、コンピュータのグラフィックインターフェースとのやりとりと少なくとも同じくらい豊富でなければならない。このとき、人は、ボックスにチェックマークを付け、キーボード上でテキストを入力し、アイコンをダブルクリックし、またはアプリケーションのウィンドウ内の十字をクリックするのと同じやり方で選択肢を選択し、単語を綴る、またはアプリケーションを実行するまたはアプリケーションを離れることができることを望む。人は、人間化され自然なユーザの為のインターフェースを望むからこそこれらの既存要素を単純に複製したいとも思わない。したがって自律ヒューマノイドロボットに適合化されたユーザインターフェース要素を見つける必要がある。

20

【 0 0 8 3 】

これらの要素はまた、ヒューマノイドロボット行動の作成者にとって容易にパラメータ化可能でなければならず、またユーザの言語に容易に適合できるようにしなければならない。既存の自律ロボットは音声認識などの簡単な人間-ロボットインタフェースを導入することができるが、従来技術では、ユーザと開発者のいずれも、地域化される（多言語使用を可能にする）とともに失敗を管理するいかなる多様式（multi modal）ユーザインターフェース要素を提供されなかった。実際、今日、多重センサ捕捉および処理能力と移動運動能力とその四肢の多数の自由度とを備えた合理的な大きさと価格のヒューマノイドロボット内に埋め込むことが可能な音声認識のタイプは、ロボット上に搭載可能なコンピュータ資源と電気エネルギー資源により必然的に制限される。これらの資源は実際には、信号の捕捉の安全性と信頼性と動きの実行に必要な制御とを保証できるようにする処理動作に対する優先度により必然的に割り当てられる。したがって音声認識のこの状況における避けられない不完全性の最良の補正を可能にする人間-ロボットインタフェース要素を提供することと、特に、ロボットがユーザから受け取るメッセージに対しロボットにより与えられる解釈についての疑問と収束する対話シーケンスの枠組み内で生ずるロボットにより返される質問とを解決するためのメカニズムによりユーザに良好な相互作用を提供すること、が必要である。

40

【 0 0 8 4 】

ロボットは人間的基準（human references）、すなわち、人間が同じ状況においてとるであろう身振りと行動を返さないのので、人間は自然なやり方でロボットに話しかけないということにも注意する必要があるだろう。相互作用はとりわけ、ロボッ

50

トが人間の方向を見なければ(人間-人間相互作用における習慣的相互作用が無ければ)自然なものとはならない。さらに、人間コミュニケーションとの対比では、多機能ヒューマノイドロボットに搭載されるコンピュータ資源に適合する音声認識のタイプはそれ自体は幾人かのユーザとの相互作用の効果的管理を可能にしない。さらに、たいていのロボットは自然言語をほとんどまたは全く使用しない。音声合成は通常、ロボットが読むことになるロボット用に考案された話または人間により書かれた電子メールであったとしても、人間により予め書かれた語句によりプログラムされる。したがって音声合成は、人間-ロボット相互作用を人間-人間相互作用に可能な限り近付けるようにする要素を欠く。従来技術の人間-ロボットインタフェースは、人間-人間自然相互作用を模擬し、この相互作用の成功に貢献できるようにする十分な多様式(multi-modality)または相互作用コードを有していない。さらに、ユーザにより既に取得された知識であってユーザが日々使用する知識でもインタフェースが呼び出せば、経験は、はるかに容易となり、ユーザのその部分について少し学習する必要があるだけとなる。したがって、仮想世界の部屋全体にわたって眼を巡らすことが、コンピュータキーボード上の矢印を押すことによるよりも頭を動かすことにより仮想現実ヘルメットによりいっそう本能的に行われることになる。

10

【0085】

本発明の解決策は、ソフトウェアとハードウェアを組み合わせた、自律ヒューマノイドロボットに適合化されたユーザインタフェース要素を提案する。上述のように使用される用語である、GUI要素をロボットの行動に置き換えることにより、BUI要素(ビヘイビアユーザインタフェース要素: Behavior User Interface Elements)、本明細書ではより一般的かつ単純にUI Elementsと呼ぶことができる、を定義する。このようなUI Elementsは、例えば、次のような行為を簡単なやり方でコード化するように定義されることができる。

20

ーロボットの頭の3つの触覚センサを同時に叩くことにより即座にアプリケーションを中止する。

ー音声認識を使用することによりロボットに尋問する。

ーロボットの触覚センサのうちの1つを叩くことによりアプリケーションの次の工程に移動する。

【0086】

30

したがってこれらの単純な要素は、ロボットのすべての行動とアプリケーションに利用可能となるようにまたは所与のプロジェクトの特定の資源を装って生成されるように一般的ライブラリに埋め込んでもよい真の相互作用コードである。

【0087】

本発明のUI Elementsは、行動開発者により使用され容易にパラメータ化することができる要素である。行動をプログラムするための基本GUI要素となるのは主にChoregrapheボックスである。特に、これらのボックスのいくつかは、グラフィックインタフェース部品を開発するためのQt(商標)環境により生成されるWidgetライブラリを使用することによりC++でコード化されたChoregrapheプラグインを含む。

40

【0088】

本発明の実施を可能にする機能アーキテクチャの簡略図を図5に表す。

【0089】

ロボットが対話者とメッセージを交換する通信チャネルの入力/出力の制御モジュール510は、図1の中央ユニット120内に取り付けられるかまたはそれに関連して取り付けられる。このモジュールは、ロボットが備える特定の通信チャネルの送信/受信前処理手段を物理的または論理的に含む。

【0090】

これに限定されないが、それぞれが受信チャネルと送信チャネルを有する3つのタイプのメッセージ通信チャネルを同図に表した。

50

【 0 0 9 1 】

タイプ1の受信チャネル521は、人間の聴覚に対応し、ロボットが音声信号（好ましくは意味論的内容を含む音声メール）を取得できるようにする。したがってロボットは図2aに表されたマイクロホン210aを備えてもよい。このチャネルの出力は通常、音声認識アルゴリズムを実行する特定の信号処理プロセッサにより前処理される。これらのアルゴリズムは、多少複雑であるかもしれない、そして使用される環境（周囲騒音、複数の話し手など）とおおよそ完全な特定の学習の実施とに応じて変わり得る有効性を有するかもしれない。しかしながらすべての構成において、認識誤りは避けられない。

【 0 0 9 2 】

タイプ1の送信チャネル531は、人間の話し言葉に対応しており、ロボットが例えば図2bに表された拡声器210bを手段として話すことができる（すなわち意味論的内容を含む音声メールを読み上げることができる）ようにする。言語、音質、リズム、口調は、状況の関数としてそして感情を表現するために変えられてもよい。しかしながらこれらの音はまた、ピープ音、予め記録された音楽であってよい。例えばモールスシーケンスのピープ音、予め設定されたコードに従う音楽もまた、意味論的内容を有することができる。と理解される。

【 0 0 9 3 】

タイプ2の受信チャネル522は、人間の視覚に対応しており、ロボットがその環境をマッピングできるようにし、アクセス可能なメモリ内に格納されていれば認識できる画像を取得できるようにする。したがって、ロボットは例えば図2aに表されたCMOSカメラ220aを備えてもよい。カメラの1つは遠方視野に専用化され、他は近傍視野に専用化されることが好ましい。有利には、画像認識アルゴリズムはロボットの対話者の顔を検出または顔を実際に認識できるように適合化される。繰り返すが、認識性能が何であれ、不確実性または誤りは避けられない。画像認識はまた、その意味がコード化により定義され得るVDU上でロボットに提示される番号数字または商標などの単純な形に適用することもできる。

【 0 0 9 4 】

タイプ2の送信チャネル532は、直接的な人間の等価物の無い人工的チャネルである。このチャネルは、ロボットの体内に埋め込まれたLEDにより生成される光信号の発射を可能にする。数多くのLEDが、特に眼、耳、胴、足上に設けられてもよい。これらは、異なる色を有し、可変周波数明滅能力を備えることができる。このチャネルは、メッセージを送る単純で強力な手段をロボットに備えさせる。特に、特定コードがユーザにより定義されプログラムされてもよい。

【 0 0 9 5 】

タイプ3の受信チャネル523は人間らしさと等価なチャネルである。但しこのチャネルはその触覚領域が制限される。触覚領域は例えば、図2aに表されたセンサ230aなどの触覚センサに集中する。ロボットの対話者は、ロボットに2進（行為の確認）またはさらに複雑なタイプのメッセージを伝えるために触覚センサを作動する。このチャネルにより受信される情報は実際、ユーザにより定義される単一的コード（罰と褒美の意味をそれぞれ有する、舌打ちとなでること）またはモールス型の連続的コードのいずれかに対応することができる。このタイプの通信チャネルを定義するために特定の触覚センサが絶対的に必要ということではない。本明細書のさらに後で説明されるようにメッセージセンサがロボットの腕および／または前腕の位置により表される連続的アナログセンサであって前記位置が対話者によりロボットに伝達される数値を表す同じタイプのチャネルを、それが対話者の接触行為を受信する限り、定義してもよい。実際、いつでも、ロボットはその関節の角度位置を知っており、したがって移動の意味が予め定義されていれば対話者の行為による移動に起因する関節の角度位置の変化をメッセージとして解釈する方法を知っている。一肢（例えば、前腕）の単純な接触もまた、ロボットの関節の角度位置のセンサにより識別することができる。衝撃または持ち上げなどのより急峻な動きは、ロボットの慣性ユニットと足の裏センサ（FSR）とによりそれぞれ検出されてもよい。

10

20

30

40

50

【 0 0 9 6 】

タイプ3のタイプの送信チャネル533は人間の身振りと等価である。頭は、ヨー角により測定される方位角の変位とピッチ角により測定される仰角の変位との2つの自由度を与られてもよい。これらの2つの動きは伝統的に承認(ピッチ)または拒絶(ヨー)のメッセージを定義する。これらはまた、ロボットが会話状態にある対話者の方にその視線を向けることができるようにする。肩、肘、手首の関節はピッチ、ロール(ロールまたは右/左ねじれ)、ヨー、ヨーの自由度を備えてもよい。手は開く能力と閉じる能力を備えてもよい。これらの関節の動きの組み合わせにより、このチャネルを通しロボットの対話者に伝達されるメッセージの内容を定義できるようにする。

【 0 0 9 7 】

他のメッセージ通信チャネル(図には表されない)が存在する、またはそれらが定義されてもよい。特に、ロボットは、赤外線、ブルートゥース、またはWi-Fiリンクにより信号を送受信することができる。したがって、特に、この目的のためにプログラムされたリモコン、例えばApple(商標)のiPhone(商標)またはモーションキャプチャおよび/または測位機能を有する別の電話を使用することにより、対話者はこのチャネルを介しロボットにメッセージを送信することが可能である。

【 0 0 9 8 】

同様に、ロボットはこれらの通信ポートを介し別のロボットにメッセージを送ることができる。

【 0 0 9 9 】

本発明によると、メッセージ通信チャネルは、異なるタイプのチャネルをハイブリッド型のチャネルにマージすることにより定義されてもよい。したがって音声認識を備えた音声チャネルと画像認識を備えた可視チャネルの出力は、その出力がデータマージ処理により改良されることになる新しいチャネルを生成するように組み合わせられてもよい。このチャネルからの出力に関する出力は、別個に採取された2つの出力より高い信頼水準のアプリオリである。

【 0 1 0 0 】

ロボットの2人の対話者541と542もまた図5に表される。当然、ただ1人または3人以上の対話者が本発明の実施シナリオでは可能である。さらに、メッセージの交換に必要な可聴および/または可視信号を送信できるようにするデータリンクによりロボットが置かれた部屋にリンクされるという条件で、対話者はロボットからいくらか離れて置かれてもよい。当然、この場合、物理的接触を必要とするタイプ3の通信チャネルの利用は不可能となる。

【 0 1 0 1 】

ロボットのその対話者に対するおよびロボットのその環境に対する相対位置はまた、人間/ロボット対話の性質を特徴付けるようにおよび随意的にその進行を修正するように特定のセンサ(対話者の場所に関連する音声認識; 画像認識; 超音波センサ等)により測定され、例えば体積、トーン、または表情の解析により解釈され相互参照されることができる。したがって、近づき大声で話す対話者は、ロボットにより脅威と見なされ、関連する身体言語により、または相互作用の実際の修正または遮断により様々な防衛行動をトリガすることがある。

【 0 1 0 2 】

これらの様々な通信チャネルの入力/出力の論理的制御はモジュール510により行われる。

【 0 1 0 3 】

後者は、本明細書のさらに後で説明されるように同じ第1のタイプの送信チャネル上に送信されたメッセージを通して第1のタイプの受信チャネル(例えば、音声チャネル)の入力についての疑問を同一時間に解決できるようにする。上記疑問解決行為は恐らく、同じ第1のタイプのチャネルまたは第2のタイプの受信チャネル(例えば、触覚チャネル)上で対話者によりそれに応じて行われる。第1のタイプのチャネル(例えば、音声チャネ

10

20

30

40

50

ル) 上で受信されたメッセージについての疑問の解決を要求するメッセージはまた、第2のタイプのチャネル(例えば、LED送信による可視チャネル)上で送信することができ、対話者の疑問解決行為は第3のタイプの受信チャネル(例えば、触覚チャネル)上で行われなければならない。これらの組み合わせは純粋に非限定的例示として与えられ、様々な組み合わせが可能性である。

【0104】

通信チャネル510の入力/出力の制御モジュールもまた、メッセージの入力をより簡単に組み合わせるために使用することができ、この組み合わせによりロボットの「心」の中の疑問のいかなる可能性も実質的に除去できるようにする。

【0105】

受信チャネルにより受信された入力と受信チャネルにより送信された出力とを組み合わせるための機能のプログラミングは、BUI Element sを使用することにより単純なやり方で実行されてもよい。

【0106】

我々は、選択タイプの制御ボックスまたは選択ボックスからなるBUI Element のタイプについて後で説明するものとする。後者は、閉じたりストから選択を行う方法を表す。それは特に、対話の枠組み内で、限定された数の単語と語句の認識に適合化され、ロボットはユーザの選択を聴く前に質問を投げかけることができる。

【0107】

我々は、異なるタイプの選択ボックスとは別個のタイプのBUI Element について以下に説明する。

【0108】

我々は、整数を選択する例によりこの様式を示す。この要素に関し、ロボットは例えばタイプ1の送信チャネル531上で、ユーザに利用可能な最小数と最大数について述べ、一方の腕を対話者の方へ伸ばす。ロボットは弱くサーボ制御される。この腕は、図5のタイプ3の受信チャネル523を構成することになる。腕の低い位置は最小桁に関連付けられ、高い位置は最大桁に関連付けられる。したがってユーザは桁を選択するためのカーソルとしてロボットの腕を利用する。ロボットは、肩のピッチ関節(Shoulder Pitch)上で利用可能な触覚のおかげで腕の位置を知る。この相互作用を補強するために、ロボットはユーザがロボットの腕を動かしている間にその手を見る。位置の各変化によって、ロボットは選択された桁を言うことができる。ユーザは、タイプ3の別の受信チャネル523を使用することによりロボットの頭の中央の触覚センサに触れることにより、選択を検証することができる。特にセンサの精度に対し余りにも多過ぎる数の桁がある場合、一方の腕により粗調整がなされ、そして第2の腕がより正確に選択するようにしてもよい。表情の順序付きリストを数字により表してもよい。このとき上記手順は、ロボットにより告げられたドロップダウンメニューからの選択の様式になる。

【0109】

桁を選択できるようにする変形形態は、触覚センサだけを使用することからなる。例えば、

- － 前方センサを叩くことで、桁のリストにおいて1段階下ることができるようにし、
- － 後方センサを叩くことで、桁のリストを上ることができるようにし、
- － 前方または後方センサを押したままにすることで、桁のリスト内のスクロールを加速できるようにし、
- － 選択は中央センサに触れることにより行われるだろう。

【0110】

本発明の使用のシナリオの関数として、可能な組み合わせを大幅に変えることが可能であることが理解される。

【0111】

図6は、本発明のいくつかの実施形態のロボットの行動/相互作用のプログラミング論理チャートである。

10

20

30

40

50

【 0 1 1 2 】

同図により示された例は、ロボットに単語のリストからの選択を提示する対話者とロボットが対話するシナリオである（例えば、なぞなぞのゲームの場合の）。このシナリオでは、タイプ1の受信チャンネル、タイプ3の受信チャンネル、タイプ1の送信チャンネルが使用される。

【 0 1 1 3 】

図中のコード610により表される行為は、ロボットの対話者の行為：すなわち、例えばロボットにより既に告げられたリストからユーザにより告げられた選択；タイムアウト（すなわち、無選択）；このリスト内の1つまたは複数の単語の理解の確認要求に対する回答「はい／いいえ」である。

10

【 0 1 1 4 】

図中のコード620により表される行為は、コード630により表される内部変数の状態の関数として起動されるロボットの行為である。これらの内部変数の意味は以下の通りである。

- － r : 選択のリストの中からユーザにより告げられた単語のロボットによる認識の確率の比、
- － f : 認識失敗の総数、
- － t : タイムアウト（または、所定の時間後の対話者による選択が無い）の数、
- － S 1 : 認識確率比の閾値1、
- － S 2 : 認識確率比の閾値2、
- － t m a x : 可能なタイムアウトの最大数、
- － f m a x : 可能な失敗の最大数。

20

【 0 1 1 5 】

タイムアウトが処理される一般的な方法は、日々の人間生活の単純原理について投げかけられる問題へ適応「一言も言わない人は誰であろうと．．．ということに同意する」に相当する。

【 0 1 1 6 】

同図に表される処理動作の一般的論理について以下に説明する。

【 0 1 1 7 】

N A O はユーザ／対話者に耳を傾け、変数 f と t はゼロに初期化される。所定のタイムアウト時間が経過したことを対話者が認めればタイムアウトカウンタはインクリメントされ、タイムアウトの最大数に達すれば相互作用ループが遮断される。

30

【 0 1 1 8 】

このアプリケーションは、開始される際にプレーヤの数を確かめるためにゲームにおいて、ユーザにより行われるロボットの尋問などの特定の行為が、または頭上の触覚センサの1つを押すことにより、このアプリケーションをトリガする決定論的状况において、または人間の存在の検出、時刻、またはより一般的には一日の事象のログ（ロボットにより格納される）などのパラメータの関数としてこのアプリケーションをトリガする人工知能の状况において、のいずれかの状况における行動で初期化されてもよい。例えば、ユーザがこのアプリケーションを呼び出し中であることを検知すると、ロボットはユーザがその何を望むかおよびなぜそれを呼び出したかを知ることができるようにするアプリケーションをトリガする。別の状況では、人間の存在を検出すると、ロボット自体は、ユーザがプレーすることを大いに望み、かつユーザがプレーしてから長い時間経っているゲームを提案するためにアプリケーションをトリガすることができるようになる。

40

【 0 1 1 9 】

対話者がタイムアウトの終了前に選択を告げると、測定された認識確率比 r は期待認識確率比の閾値 S 1、S 2（S 1 < S 2）と比較される。これらを判定するやり方については後で説明する。

【 0 1 2 0 】

$r \leq S 1$ ならば、この単語の認識は失敗と考えられる。失敗カウンタがインクリメント

50

される。 f_{max} に達すれば、単語は未認識であると明確に宣言され、相互作用は遮断される。 f_{max} に達しなければ、以下の3つの場合の準備が図に示すようになされてもよい。

－ 第1番目の失敗($f = 1$)では、ロボットは「分かりませんでした」とその対話者に示し、選択のリストの繰り返しからなる機能「`activateHelpWhenFailure`」の1つを起動する。

－ 第2番目の失敗($f = 2$)では、ロボットはまた「分かりませんでした」と示し、対話者に選択のリストを提供することと、対話者にその触覚センサを使用するように依頼し、それをどのように使用するかを示すことからなる機能「`activateHelpWhenFailure`」の別のものを起動する。

10

－ それを超える($3 \leq f < f_{max}$)失敗では、ロボットは、効果的な会話の条件が満たされていないということを示す語句、例えば、通常は前記対話者を促して会話を終了させることになる「雑音が多すぎる」などを告げることができる。

【0121】

$S1 < r \leq S2$ であれば、ロボットは実際に聞いたものに関し疑問を抱き、同図に表された手順に従って、認識したと考える単語または表現を発声し、その対話者に「これは正しいですか?」と尋ねることにより疑問を解消する行為を取ることができる。対話者が「はい」と答えるか、またはタイムアウトの終了時に答えなければ、ロボットは回答が正しいと考える。対話者が「いいえ」と答えれば、失敗カウンタはインクリメントされ、 f_{max} に達すると、ロボットは、分からなかった、と明確に示し、相互作用は停止する。 f_{max} に達しなければ、

20

－ 第1回目の失敗($f = 1$)では、ロボットは、選択のリストの繰り返しからなる機能「`activateHelpWhenFailure`」の1つを起動することができる。

－ 第2回目の失敗($f = 2$)では、その対話者に選択のリストを提供することと、その対話者にその触覚センサを使用するように依頼し、それをどのように使用するかを示すことからなる機能「`activateHelpWhenFailure`」の別のものを起動する。

－ 第3回目から f_{max} までの失敗では、対話者は認識の確率比が改善するまで選択を繰り返さなければならない。

【0122】

30

このようにして、音声認識の不完全性を著しく緩和することと、ロボットとその対話者との会話における流動性を改善することが可能である。

【0123】

図7a、7b、7cは、本発明のいくつかの実施形態における多様式インターフェースの相互作用の論理的および時間的組み合わせを示すタイムチャートを表す。

【0124】

これらの図は、図6のチャートに表されたタイプの相互作用をプログラムできるようにする選択ボックスの図である。

【0125】

選択ボックスは、項目410下で示されたものなどのボックスであるが、これらは自然な対話のための特定の行動の特に効果的なプログラミングを可能にする特定タイプのものである。

40

【0126】

これらの図面内の符号の意味は以下の通りである。

－ 図7aでは、

－ 710aはロボットまたはその対話者の行為／スピーチを示す。

－ 720aは触覚センサを示す。

－ 740aは認識ビーブ音を示す。

－ 750aは回転するアニメ化された位置のロボットの顔のLEDを示す。

－ 751aはロボットの固定位置の顔のLEDを示す。

50

- － 7 6 0 a はロボットの顔のLEDの点滅を示す（受信されたメッセージのロボットによる理解の関数として様々な色であってよい）。
- － 7 7 0 a はタイムアウト機能を示す。
- － 7 8 0 a は選択ボックスの出力を示す。
- － 7 9 0 a は機能「触覚センサメニューに行く」（図7 b）を示す。
- － 7 A 0 は機能「選択のソートに行く」（図7 c）を示す。
- － 7 B 0 は機能「音声認識メニューに行く」を示す。
- － R 1、R 2、R 3 は、ロボットが曖昧さ無しに理解した場合、ロボットが理解するがいくつかの疑問を抱く場合、ロボットが全く理解しない場合をそれぞれ示す。
- － 図7 c では、7 1 0 c は機能「前ページのメニューに戻る」を示す。

10

【0127】

選択ボックス内でプログラムされた処理動作の一般的論理は既に説明したものと同一である。ここで説明される追加要素は以下のとおりである。

－ ロボットの顔のLEDs 7 5 0 a の利用、随意的には一問一答を強調するためにLED点滅の利用。LEDは、ロボットがスピーチを検出し解析中であることを示すために固定位置7 5 1 a にある。

－ 認識の準備ができていたを示すためにロボットにより発せられる可聴「ピープ音」の利用。実際、処理能力と電源の制限のために、そしてまた認識時のうるさを回避するために、後者は音声合成と同時に活性化されない。したがってロボットにより対話者に投げかけられた質問は後者によりあまりにも早く答えられる必要はない。「ピープ音」は、答え始めるためにピー音を発する。

20

－ このやりとりと以前のやりとりの間にこのロボットの履歴ログとユーザのその経験とに応じていくつかの水準のヘルプを利用する可能性。

－ プログラミングを容易にするためにいくつかのメニューをナビゲートする可能性。

【0128】

次に説明する図面は、上の図4 に対する注釈として述べたChoregrapheソフトウェアの選択ボックス部品が、表された例ではタイプ1（音声交換）の受信チャンネルと送信チャンネルを使用することによりNAOロボットと対話者間の単純または複雑な相互作用をプログラムするために利用される画面ショットである。

【0129】

30

図8 a、8 b、8 c、8 d、8 e は、本発明の一実施形態における対話言語を変更するための2 進選択および選択肢によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。

【0130】

図9 a、9 b、9 c、9 d、9 e は、本発明の一実施形態における対話言語を変更するためのリストと選択肢からの選択によりヒューマノイドロボットとの対話をプログラムできるようにする一連の画面を表す。

【0131】

図1 0 a、1 0 b、1 0 c、1 0 d は、本発明の一実施形態における選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。

40

【0132】

図1 1 aと図1 1 b は、本発明の一実施形態における、選択のリストの選択肢を置換または補足することができるようにし、いくつかの選択肢同士を比較する新しい音声認識試験を実行できるようにする一連の画面を表す。

【0133】

図1 2 a、1 2 b、1 2 c、1 2 d は、本発明の一実施形態における、質問のものと異なる言語で選択のリストのいくつかの選択肢同士を比較する音声認識試験を実行できるようにする一連の画面を表す。

【0134】

図1 3 a、1 3 b、1 3 c、1 3 d は、本発明の一実施形態における、選択のリスト

50

のいくつかの選択肢同士を比較する音声認識試験の閾値を検証／修正できるようにする一連の画面を表す。

【0135】

通常、選択ボックスは、ユーザが選択の事前定義セットの中から回答を選択できるようにする。選択ボックスは、開発者が直観的かつ可読なやり方で一組の可能な選択を書くことができるようにする表型の部品を呼び出す。開発者が予め知らなければ、選択のリストもまたボックスへの入力として入力することができる。したがって例えば、ユーザのメールを管理するアプリケーションの場合、ロボットはユーザに、別のファイル内に格納されたユーザのアドレス帳から連絡先を選択させることができる。これらの `UI Elements` は高度にパラメータ化可能なツールである。したがって音声認識および／または合成を使用する `UI Elements` は地域化される。例えば選択ボックスは仏語と英語で編集可能である。そのプログラミングのためのグラフィックインターフェースレベルでは、ボックスを編集するための言語を変更するために使用される `Widget Qt` (商標) が `ComboBox` であってもよい。

10

【0136】

`Choregraphe` ボックスの入力(及び出力)は以下のいくつかのタイプのものであってよい。

1. 「バン音(`bang`)」: 信号が送られる。
2. 数字: 入力は整数または浮動小数点の数を取り出す。
3. 文字列: 入力は文字列を取り出す。
4. 動的。

20

【0137】

動的型の入力(各々出力)は `ALValue` を取り出す(各々出力する)。

【0138】

`ALValues` は `NAOqi` ライブラリ内に記載された一般型(特には、整数、浮動小数点、配列、ブーリアン、文字列、未初期化 `ALValue` である「`bang`」)の集合である。動的型の入力はアプリケーションの展開を極めて柔軟なやり方で管理する。特に、様式間および／または様式内確認モードの選択と支援の提示は、可能な選択の数に応じてそれらを起動するためにロボットの対話者に提供される。

30

【0139】

したがって、この入力が動的型であるという条件で、`python` 表(配列型)を `Choregraphe` ボックスに対する入力として入力することが可能である。

【0140】

本発明を実施するために使用される `Choregraphe` ソフトウェアは、ブーリアン型のボックス(チェックボックス)、文字列タイプのボックス(テキストボックス)、編集可能またはエンドユーザによらない文字列タイプの多肢選択のボックス(コンボボックス)、整数または浮動小数点浮動小数点型のボックス(スライダ)、または他のタイプのボックスのパラメータを含む。例えば、その行動またはアプリケーションにおいて選択ボックスを使用するプログラマは、ブーリアンパラメータ「検証された選択を繰り返す」(仏語では、「`Repetere le choix valide`」)のチェックマークを付けるまたはチェックマークを外す可能性を有する。これは、`NAO` がユーザにより検証された選択を系統的に繰り返すかどうかを定義するので、相互作用中の `NAO` の行動に影響を及ぼすことになる。

40

【0141】

音声認識の欠陥を緩和するために、診断ツールは音声相互作用の成功を最大化できるようにする。したがって、選択ボックスにおいて、開発者が表中の単語のリストの書き込みを完了すると、これらの単語の百分率認識(例えばロボットにより確かに認識される単語に対応する100%、ロボットが認識しない単語に対応する0%)を示すことになるこのツールを実行することができる。この診断は、音声合成により話された単語(ユーザが話すものに近いと仮定する)と音声認識により予想される単語との比較により行われる。さ

50

らに、選択毎に相互作用の機会と自然さとを最大化するように、いくつかの表現を定義することができる。したがって、メールを送信するようにロボットに依頼するために、開発者は、ユーザの思うままに「メールを送る」、「メッセージを送る」、「電子メールを送る」などのいくつかの語句を配置することができるようになる。ユーザは、その後、最終的に同じことを言うようにこれらの様々な表現同士間の選択肢を有することになる。

【 0 1 4 2 】

本発明の解決策はまた、幾人かのユーザの存在を管理しない音声認識の問題を解決できるようにする。人間は、数人と話をする際にはコミュニケーションが困難であると認識しているので、一人ずつ話すことにより適応する。この状況は、よく知られたフランスの「t u」形式のロボットによる利用などの明確に単一ユーザ相互作用コードの存在により容易にされる。

10

【 0 1 4 3 】

欠点のある音声認識は、人間-ロボットインタフェースが特に失敗の状況を最もよく管理し、ユーザに正しい時に話させ（これは相互作用コードに関わる）、対話の代替であってより効果的な解決策を利用できるようにさせなければならないということを必要とする。

【 0 1 4 4 】

本発明の枠組み内では、音声診断機能がこのタイプの問題を解決できるようにする。この機能は、試験対象の単語を音声合成ソフトウェアである、text-to-speechにより発音させることにより実行する。次に、この試験対象の単語は音声認識により解析される。より正確には、同じ単語は例えば3回発音され、毎回、単語を発音するやり方の代表的サンプルを有するように音声とそのピッチの速度を変化させる。次に、音声認識により返された3つの認識率が平均化される。単語の推定百分率認識はこの値である。音声診断の2つの可能なモードがある。

20

【 0 1 4 5 】

その「一斉(T o g e t h e r)」モードは次のように動作する。選択ボックス内に記録されたすべての単語は音声認識により聴取され、次にN A O が、他で説明されるように推定認識率を計算する。

【 0 1 4 6 】

「一つずつ(O n e b y O n e)」モードは次のように動作する。他のライン上の他の可能な選択肢だけでなく所与のラインに関し、解析対象の単語が音声認識により聴取されるが、それと同じライン上に置かれたその別の単語は聴取されない。この診断の利点は、2つの「同義語」が互いに似ている場合、例えば「c o u c o u !」と「c o u c o u t o i !」、推定認識率は「一斉」モード（推定認識率は、これらが音声認識によりしばしば混同されるので極めて悪くなるであろう）ほど低くはならない。実際、2つの同義語がロボットにより混同されたとしても深刻ではない。

30

【 0 1 4 7 】

診断が各ラインに対して行われると、同義語が推定認識率の降順で配置され、最良の同義語の認識率はラインの終わりに記録される。

【 0 1 4 8 】

40

したがって選択ボックスは、ロボットがそれを正しく認識または解釈したかが定かでないときにユーザに回答を確認することを依頼するようにプログラムされる。このメカニズムは、聴力障害を有する人間または理解を困難にする環境に陥った人間により使用されるものと同一である。ロボットは、ユーザの回答の理解のレベルに応じて異なる反応を有することになる。次に、いくつかの閾値（例えば、図5に対する注釈として定義された閾値S 1 とS 2）は、認識ソフトウェアにより計算される認識信頼度の関数として固定される。例えば、第1の認識閾値S 1 に達しないとき、ロボットはプレーヤに回答を繰り返すように依頼し、第1の閾値S 1 に達したが第2の高い認識閾値S 2 に達しないとき、ロボットは、それに対する回答が疑問を解消できるようにする質問を投げかけることになる。ロボットはまた、ユーザがロボットに対し正しく答えるようにヘルプを与えることができる

50

。すなわちロボットは可能な選択のリストを与え、それとの相互作用手段を示すことができ、投げかけられた質問があればこれを繰り返すことができる。相互作用コードはまた、音声認識の欠陥を緩和するのに極めて役立つ。実際、音声認識は、ロボットが話している間はロボットに話しかけることを可能にせず、音声認識の実行とそれが実際に活性化される瞬間との間の遅延はかなり長い。したがって、音声認識が実行されると可聴コードが働き、話すことができるということをユーザに示す。その後、かなり直観的な可視コードと回転する耳のLEDが、ユーザに、ロボットは聴いているということを知らせる。音声認識を使用するUI Elementsはまた、この音声認識に代替手段を提供し、繰り返される理解の問題の場合（これは例えば極めて雑音の多い環境によるかもしれない）にもユーザが成功裡にコミュニケーションできるようにする。これらの代替手段は触知可能、可聴、可視的なものであってよい。例えば、選択ボックスはユーザが触覚センサを使用することにより回答を選択できるようにし、前方センサを押すことで選択のリスト内を進むことができるようにし（このとき、ロボットは各選択を告げる）、後方センサはこのリスト内を後退できるようにし、中央センサは選択を検証できるようにする。ロボットが様々な選択を告げることと、ユーザが検証したい選択を聞いたときに「OK」と言うこと、とを構想することも可能である。または、そうでなければ、確認のために、「はい」または「いいえ」と答える代わりにユーザはロボットの一方の腕を押すことができる。図5に対する注釈として定義された様々なタイプ1、2、3の通信チャネルの入力／出力の制御モジュールは、選択ボックスの様々な入力／出力間のリンクを通してこれらの組み合わせを管理するための機能を単純かつユーザフレンドリなやり方で生成できるようにする。

10

20

【0149】

通常、本発明の解決策は、インターフェースの人間化、すなわち人間－人間インターフェースの模擬形態を提案する。我々は、スピーチ（もちろん、すなわち発言された単語群）だけでなく口調と視覚的要素との3つの主要素が2人の人間同士の直接コミュニケーション中に作用し始めるということを知っている。その証しとして、それらの進展を通して、書くことまたはインスタントメッセージなどの間接コミュニケーション手段を観測することにより、対話中の情報不足が一般的な法則として、直接コミュニケーションの代用物と句読点または最近の顔文字などの代用物の追加によりどのようにして緩和されることができるということを極めて明確に理解することが可能である。すべての場合において、今日の大きな技術的進歩にもかかわらず、これらの基本要素が全体として人間－ロボットコミュニケーションのために置き換えられることは依然として難しい。但し、対話の演出を改良する人工的代用物を見つけることは可能である。ロボットの音声合成と音声認識はスピーチの等価物を可能にする。したがってこれらは人間とのコミュニケーションの柱である。ヒューマノイドロボットはさらに、対話の視覚的要素の大部分、すなわち身振り顔の表情、を表現することができるという利点を有する。実際、その擬人的ボディにより、その移動は車輪上のロボットほど容易でないが、その身振りは人間行動に、より簡単に基づくことができるので、人間の動きとして容易に解釈することができる。このとき、コミュニケーションはより自然に実行される。

30

【0150】

それにもかかわらず口調と顔の表情は、固定された顔と口調とを有するロボットには欠けている。但し、これらの2つの要素は、これらの要素を伝える他の機能とコードにより補償される。これらは、ユーザによる多少長い学習を必要とする。このとき、目的は、この学習を可能な限り短くし、したがってユーザが既に知っているものに対し可能な限り一貫性がありそれに近いコードにすることである。

40

【0151】

Ben Shneidermanの著作である、「Designing the User Interface: Strategies for Effective Human-Computer Interaction」(1997年発行: <http://www.cs.umd.edu/hcil/pubs/books/dtui.shtml>)の中で述べられ、通常はグラフィックインターフェースに適用される人間工学の基

50

本法則を適合化することにより、単純でかつ一貫性のあるコードとしたがって自然でかつ流動的相互作用を実現する。これらの法則は次の原理：コードとインターフェース要素との一貫性、上級ユーザのためのショートカットの存在、行われる行為に関する即時復帰の存在、対話の明確な終了、誤りの簡単な管理、フィードバックの可能性、ユーザは相互作用中は自分自身をマスターと感ずるにちがいないこと、そして最後に、ユーザの短期記憶の軽い刺激について述べている。

【 0 1 5 2 】

音声認識および合成は、特に自然言語の欠落と限定数の単語だけを認識できるようにする単独ユーザ認識とにより、限定的である。本発明の解決策は、十分に自然な人間-ロボット相互作用を提供するように、ロボットによる自然言語の不使用の問題を解決する。既に、最良の使用はロボットの音声合成からなる。特に、音声合成および／または認識を使用するロボットのUI Elementsのほとんどは地域化される。したがって仏語を話す（英語を話す）ユーザは仏語（英語）でロボットと対話することができこれにより相互作用の成功を最大化する。その後、最良の使用は、ロボットの反応性を改良するとともに人間-ロボットコミュニケーションの成功を容易にするために、タイミングと相互作用コードからなる。したがって、選択ボックスは、ユーザからの回答の待ち時間のようないくつかのパラメータを提案する。したがってユーザが全く答えなかったということを考慮するまでにロボットがあまり長く待たないということだけでなく、音声認識を正しい時に起動することが可能となるためにロボットが十分に長く待つことも保証される。相互作用コードは、身振り、可聴および／または可視のものであってよい。したがって音声認識の終了の可聴ピープ音は、ユーザに、ロボットがもはや聴いていないということを知らせる。

10

20

【 0 1 5 3 】

さらに、本発明の解決策では、コミュニケーションは、異なる様式のいくつかの通信チャネルの利用とロボットの一部分に関する特定の行動とによりより自然なものにされる。したがって、音の場所と顔の検出（特にその位置）との利用により、ロボットが人間の対話者の方に頭を回転できるようにする。これは別の人間に対処する際の確定事実であるように思われる。ロボットはまた、特に名前、固有の特徴（例えば、ロボットにより行われた会話と行動の履歴ログ）を利用することにより人間に対処するように、対話者識別（顔認識、音質、声紋など）を実施することができる。ロボットはまた、触覚センサをなでたかどうかに応じて行動をユーザが何と考えた（その人はその行動が好んだ）かを知ることができ、その後、例えば言語コミュニケーション中にその行動を実行することを提案する。ロボットは、状況に適合したやり方で行動しようとする。したがってロボットはアニメーションを再生し、そのLEDを使用し、音を放送することができ、これにより人間が話すときに人間がなす本能的な身振り（両手などにより話すこと）を模擬できるようにする。ロボットはまた、うなずくことができる。いくつかの検討、特にJustine Cassellの記事、「Social Dialogue With Embodied Conversational Agents」（2005年発行：<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.124.9853&rep=rep1&type=pdf>）においてなされた検討は、人々は、うなずきが無作為に生じた場合でも機械または化身がうなずくと機械または化身が会話にさらに興味を持っているようだと考える、ということを証明できるようにした。これらのあらゆる身振りの信号（頭、腕または手により同意または反対することなどの）、音声信号、相互作用の年表（chronology）、対話者の場所、またはそうでなければユーザの意図（ユーザは前進しているのか後退しているのか）の検出は、人間-ロボット相互作用を通常の人間の規範により近付けることによりそれをより自然かつ効果的にすることができる。したがってこれらはまた、制限的音声認識に関する問題のいくつかを解決する。これらは本発明の利点の1つである。図8～図13の画面ショットは、NAOロボットと対話者との対話のいくつかのケースをプログラムしてもよいやり方を説明する。

30

40

【 0 1 5 4 】

50

行動において選択ボックスを使用するためには、選択ボックスをChoregrapheのデフォルトライブラリから図(図8a)内にドラッグアンドドロップする必要がある。質問を仏語と英語で編集できるようにするローカライズテキストボックスが追加される。質問「あなたのお気に入りの動物は何ですか?」はローカライズテキストボックスのテキストプラグイン内に英語で書き込まれる(図8b)。人はまた、質問を仏語で編集することを望む。したがって、ボックスのプラグインのコンボボックスが用いられ、仏語が選択される(図8c)。このとき、質問のテキストは、編集されないときは空であるテキストプラグイン内に仏語で書き込まれる: "Quel est ton animal préféré?"(図8d)。質問は、ロボットにより正しく管理されるように選択ボックスに入力されなければならない。ローカライズテキストボックスの出力は選択ボックスのonStart入力にリンクされる(図8e)。我々は今、選択ボックスのコンボボックスにより示されるように選択を英語で編集するものとする。デフォルトの選択がボックスから取り除かれる。第1番目の選択「センザンコウ」が表の第1行に書き込まれる(図9a)。第2行目では、別の動物「蜘蛛」が提案されるが、同義語「タランチュラ」がスラッシュ「/」により蜘蛛から区切られて追加される。これは、ロボットが蜘蛛とタランチュラが同意語であると考えてることを意味する(図9b)。選択の編集は例えば「兎」と「ポニー」を追加することにより終了される。行の数は選択が追加される際に自動的に適応することに留意されたい(図9c)。コンボボックスは選択ボックスを編集するための言語を仏語に切り替えるために使用される(図9d)。英語と全く同様に、選択のリストが書き込まれ、これにより「pangolin」、「araignee/tarentule」、「lapin」、「poney」を得る(図9d)。

10

20

【0155】

しかしながら、我々はロボットが実際これらの選択を認識するかどうかを知らない。次に、我々は音声診断機能を実行したい。我々は「プラス」をクリックする(図10a)。次に、我々は評価アイコンをクリックする(図10b)。行毎に、単語は評価される(図10c)。同義語の場合、最良のものが初めに配置され、示される百分率はこのときの最良同義語の結果である。音声診断は終了し、このとき「poney」は極めて不完全に認識されている危険があることが分かる(図10d)。

【0156】

次に、我々は、単語を変え、「cheval」をその代りに入れることを決定する(図11a)。診断が再実行される。「Cheval」は82%の優れた得点を獲得し、我々はそれを保持する(図11b)。

30

【0157】

我々は英語に切り替え、単語に対し診断を英語で実行する(図12a、12b)。次に、「cheval」の翻訳として同義語「馬」が「ポニー」に加えられる(図12c)。診断が再実行され、「ポニー」より良い得点を有する「馬」が第1番目の位置に自動的に配置されることに留意されたい(図12d)。

【0158】

我々は、次に調整してもよいパラメータを編集する。我々は選択ボックスの左下のサムホイールキーをクリックする(図13a)。パラメータタ化ウィンドウが開く(図13b)。我々はブーリアンパラメータ「腕を動かす」にチェックマークを付ける(図13c)。したがってロボットは話している間、その腕を動かすことになる。我々は、この新しいパラメータ化を検証するためにOKをクリックする。

40

【0159】

行動の一般的入力ローカライズテキストボックスの入力にリンクされ、選択ボックスの出力は行動の一般的出力にリンクされる(図13d)。

【0160】

上に述べた例と同様にプログラムされたソフトウェアの例示的な動作について次に説明する。

【0161】

50

ロボットは、Choregraphe アイコン「すべてのモータのオン／オフをスレーブ化する」によってスレーブ化され、次に、姿勢ライブラリの「init pose」位置によって直立状態にされる。ロボットの言語は、そのウェブページ上に存在するパラメータによって仏語に設定される。行動は、Choregraphe のプレーアイコンによってロボット上で開始される。

【0162】

腕を動かしながら、ロボットは「あなたのお気に入りの動物は何ですか？」と尋ね、次に、傾聴音声信号(listening sound signal)を発する。聴いている間、眼は耳と同様に青色になり、頭の触覚センサは青色で明滅する。

【0163】

次に、ユーザは「イルカ」と答える。何が言われたかを解析している間、NAOの眼は黄色になる。NAOは回答を理解しなく、眼は2度赤色で点滅し、耳は2度青色で点滅する。NAOは、腕を動かしながら「分かりませんでした。あなたが回答できます：センザンコウ、蜘蛛、兎または馬。あなたのお気に入りの動物は何ですか？」と言い、聴取段階(listening phase)に戻る。

【0164】

次に、ユーザは「兎」と答える。ロボットは確信が持てないが、センザンコウと理解したと信じる。眼が緑色で一回点滅する。次に、ロボットは活発に腕を動かしながら、「私はセンザンコウと理解しましたが、正しいですか？」と言う。ユーザは「いいえ」と答える。ロボットは眼を赤色で一回点滅し、そしてその腕を動かしながら支援を開始する。「センザンコウ、蜘蛛、兎または馬？あなたは、私の触覚センサの助けを借りて回答を選択することもできます。あなたのお気に入りの動物は何ですか？」そして傾聴モード(listening mode)に戻る。次に、ユーザは前方触覚センサを押し、ロボットはその眼を青色で一回点滅し、「センザンコウ」と言う。次に、ユーザは再度押し、ロボットはその眼を青色で点滅しながら「蜘蛛」と答える。3回目に、ロボットは眼を青色で一回点滅ながら「兎」と言う。次に、ユーザは、ロボットの選択を検証するために中央の触覚センサを押し。ロボットは眼を一度緑色で点滅し、次に「兎」と繰り返し、ボックスと行動を終了する。

【0165】

以下に説明されるものなどのような、ロボットの通信チャネル間の他の相互作用が可能である。

【0166】

選択ボックスは、好ましいやり方で、ユーザの選択を認識するように触覚センサと組み合わせて音声認識を利用する。別の可能性はロボットの視覚(特に画像認識)を利用することである。これは概念認識ではなく物体認識である。瓶が示されれば、画像認識は別の商標のものではなくこの同じ瓶を認識することになる。本発明を実施できるようにするバージョンの開発ソフトウェアの可能性の1つは、このソフトウェア内にロボットのカメラリターン(camera return)を有することである。ユーザはロボットに物体を示し、Choregraphe において得られた画像を見て、画像内の興味のある対象物を意のままに特定することができる。ユーザはそれに名前を付ける。次に、ロボットは対象物を解析し、それを画像のデータベースに格納する。次に、ユーザは選択ボックスの可能な選択としてこれらの画像を利用することができる。例えば、ユーザが選択ボックスに「瓶」、「コップ」、「雑誌」などの対象物の名前を記入することを望む場合。ユーザは選択ボックスにこれらの単語を記入し、次に瓶、お気に入りのコップ、雑誌のカバーを取り上げ、それらをロボットが先に説明したように解析するようにロボットに示す。次に、選択ボックスはロボットの画像のデータベース全体を探索する。「コップ」と称する対象物が存在すれば、NAOはユーザの話を聴くと同時にそれを探し、等々、他の単語も同様である。こうして、ユーザは、ユーザの選択を聴くNAO上のこのボックスを開始する。ユーザは「瓶」と言ったがロボットは理解しない。2人のやり取りの後、ロボットは、「瓶」、「コップ」、「雑誌」がそのデータベース内にあるのでユーザにこれらを示すこ

10

20

30

40

50

とができることを説明する。ユーザは、ロボットの話の聴きながら、記録に役立つ（または同じ商標の）瓶を示すことができる。次に、ロボットはあたかも単語「瓶」を認識していたかのように振る舞う。

【 0 1 6 7 】

本発明の枠組み内では、ロボットのユーザのメッセージングアカウントを受信する／読む、書く／送信する、管理するエージェントとして機能するようにロボットをプログラムすることも可能である。このアプリケーションについて以下に説明する。

【 0 1 6 8 】

メールアプリケーションにより、N A O は特に電子メールを読む、電子メールに応答する、または連絡先に電子メールを送信するだけでなく、受信されたメールの筆者を連絡先に加える、メッセージを削除する、メッセージを未読としてマーキングする、再読する、次のまたは前のメッセージを読むこともできる。

10

【 0 1 6 9 】

このアプリケーションでは3つの選択ボックスが使用され、これを不可欠の要素とする。単語は音声診断によって選択されている。

【 0 1 7 0 】

アプリケーションが実行されると、ロボットはユーザが新しいメッセージを受信したかどうかを調べることから始める。そうならば、ロボットは最初の新しいメッセージを読み、次に、質問を伴わずに選択ボックスを実行する。そうでなければ、ロボットは、この同じ選択ボックスを次の質問を伴って実行する：「あなたは私に何をしたいのですか？」。したがって、質問を伴うまたは伴わずに選択ボックスを実行できるということが、メールアプリケーションにおいて利用される。この選択ボックスは、ユーザがN A O の可能な行為の中から選択できるようにする。これらの行為はボックスのプラグインの表に書き込まれる。「タイムアウト」選択ボックス出力は、タイムアウトの場合にN A O が次のメッセージを読むので、役立つ。このとき、パラメータ「無返答時の最大繰り返し回数」は1に設定される。ロボットは最初のタイムアウト時にこの選択ボックスから出る。さらに、ユーザによる選択後にロボットは理解したものを明確に示す特定のアニメーションまたは行為を実行するので、パラメータ「検証された選択を繰り返す」は非活性化される。ブーリアンパラメータ「頭を動かす」、「腕を動かす」、および「脚を動かす」のおかげで、ロボットはその決定に合わせたアニメーションにより活気付けられることになる。

20

30

【 0 1 7 1 】

例えば、このボックスの可能な選択は以下のものである。

- 返答を記録する／このメールに返答する／そのメールに返答する／返答する。
- 再度読む／このメールを再読する／そのメールを再読する／再読する。
- 次／次を読む／次のメールを読む。
- 前の／前のメールを読む／前のメッセージを読む。
- 未読としてマーキングする／保管する／後で再読する。
- 削除する／メールを削除する／メッセージを削除する。
- メールを書く／メールを送る／送る。
- 連絡先に加える。

40

- 出る／やめる／飛ばす／停止／止める／キャンセルする／静かにする。「出る」は選択ボックスのデフォルト選択の1つであり、ここではメールアプリケーションから出られるようにする。

【 0 1 7 2 】

ユーザが「メールを書く」選択肢を選択すれば、ユーザはまずアドレス帳から連絡先を選択しなければならない。この選択機能を行うために、「あなたは誰宛に書きたいのですか？」という質問を伴った選択ボックスが入力として使用される。選択のリストは可変である。その結果、選択ボックスの表は記入されなく、連絡先リストがそれを保存するファイルから読み出され、動的タイプの選択ボックスの「c h o i c e s L i s t」入力に送られる。この時、パラメータ「検証された選択を繰り返す」は、N a o がメッセージを誰

50

宛に送るかを実際に理解したということをユーザに示すために活性化される。

【 0 1 7 3 】

タイムアウトの場合に誰宛にもメールを送らず実際にはメールの送信をキャンセルしメインメニューに戻ることができるように、パラメータ「無返答時の最大繰り返し回数」は例えばそのデフォルト値3に設定される。同様に、アプリケーションのデフォルト選択である「出る」と言うことにより主メニューに戻ることができるようにするヘルプ機能はユーザが連絡先をもはや思い出さない場合のためのものである。この場合、例えば触覚センサによりNAOは連絡先のリストを告げる。

【 0 1 7 4 】

メールを直接送信する場合、またはそうでなければ受信メッセージに応答する場合、ロボットはユーザのメッセージを記録することになる。

【 0 1 7 5 】

メッセージを終了すると、ロボットは記録されたメッセージを再読し、次に例えば以下の様々な相互作用を提案する選択ボックスを開始する。

－それを再生する／メッセージを再生する／私のメッセージを再生する：NAOはメッセージを再読する。

－メッセージを再記録する／私のメッセージを再記録する／それを再記録する。メッセージは、最初のもので適切でなければ再記録することができる。

－それを送信しない／送信しない／メッセージを送信しない。NAOはメッセージを送らず、アプリケーションの前のレベルに戻ることになる。

－それを送る／メッセージを送る／私のメッセージを送る。NAOはメッセージを送る。

－タイムアウトの場合、メッセージが送られる。

－「出る」要求または繰り返しに伴う失敗などのタイムアウトでないボックスから「他の」出る場合、アプリケーションは前のレベルに戻る。

【 0 1 7 6 】

パラメータは主メニューの選択ボックスのものとほぼ同じであり、パラメータ「無返答時の最大繰り返し回数」は1に設定される。何回の無返答の後にロボットはタイムアウトがあったと考えるかを示すパラメータ「音声認識タイムアウト」と、「確認時の音声認識タイムアウト」は、ユーザが何も言わずにメッセージを容易に送ることができるように例えばデフォルト設定の6秒の代わりに4秒に設定することができる。

【 0 1 7 7 】

選択ボックスはまた、ボックスの使用の継続時間全体にわたって一定であるパラメータにより静的なやり方で構成することができる。しかしながら質問を自動的に生成するシステムの利用の枠組み内では、パラメータは自動的に調整されてもよい。例えば、As An Angel社により開発されたものなどの会話型エージェントの利用の枠組み内では、前記エージェントは、それが自動的に生成することになる、質問-回答の機能として選択ボックスを構成することができる。

【 0 1 7 8 】

特に本発明の実施を容易にするように他の改良が、Choregrapheビヘイビア開発ソフトウェアに対しなされた。以下にその説明を行う。

【 0 1 7 9 】

Choregrapheボックスは、支援されるプログラミング言語の1つのスクリプトによって実施される。このボックスが、繰り返し回数、ロボットにより利用される言語、ロボットが話さなければならないテキストなどのいくつかのパラメータ化可能な態様を有する場合、これらの情報の項目はボックスのスクリプト中に直接組み込まれる。ボックスのパラメータの修正が望まれる場合、例えばそれを違ったやり方で使用するためにそれを複製した後、その行動を変えるためにはボックスのスクリプトを修正する必要がある。これは、使用されるスクリプト言語の完全な知識の無いユーザが実行することを望むありふれた操作であって、Choregrapheユーザの生産性を改良するためのもので

10

20

30

40

50

あるので、ボックススクリプトを構成することができるよう特別のインターフェースが開発された。この機能には2つの態様がある。

【0180】

Choregraphe インターフェースでは、ユーザは、ボックスの入力と出力を生成することができるのと同じやり方で、ボックスの属性を編集するためのウィンドウ内に「ボックスパラメータ」を生成する可能性を有する。各「ボックスパラメータ」は、名前、説明、タイプ（ブーリアン、整数、浮動小数点、文字列の中から）を有する。各「ボックスパラメータ」はタイプの関数としてデフォルト値などの追加属性を有することができる。最後に、「ボックスパラメータ」は親ボックスから継承するものとして定義されてもよく、これは、値が決定されるやり方に影響を与えることになる。「ボックスパラメータ」が定義されると、ボックスは、その左下角の追加可視指標によりそのチャート内に表示される。ユーザがこのアイコンをクリックすると、「ボックスパラメータ」編集ダイアログが開き、ユーザは「ボックスパラメータ」の属性内に定義される随意的制約条件の枠組み内で、各「ボックスパラメータ」に関連する値を定義することができる。

10

【0181】

ボックスのスクリプトでは、ボックスの著者は、今後、引き数として「ボックスパラメータ」と称するいくつかの機能の助けを借りて「ボックスパラメータ」にアクセスすることができる。ボックスの著者は「ボックスパラメータ」の現在値を調べて、それを変更することができる。また、ボックスの著者は、Choregraphe 内に現れないがボックスのスクリプト内の一時記憶として機能することができる動的「ボックスパラメータ」を生成することができる。パラメータの現在値は、親ボックスから継承したものか否かマーキングされることに依存する。それでない場合（デフォルトの場合）、「ボックスパラメータ」はボックスに固有である。ボックスのスクリプトがそれを調べると、その現在値が単純に返される。パラメータの現在値が、継承するとしてマーキングされると、値の読み込み中に、同じ名前の「ボックスパラメータ」を含む親ボックスが見つかるまでボックスチャートの階層中の逆追跡（backtracking up）が発生する。何も見つからなければ現在のボックスの現在値が使用される。

20

【0182】

さらに、ロボットがそのカメラの視界に入る対象物を認識できるようにするソフトウェアモジュールが、ロボットに利用可能である。但し、認識される対象物は、最初に学習段階において学ばなければならない。この学習はChoregraphe 内の特定インターフェースの助けを借りて実行される。

30

【0183】

このインターフェースは、ロボットのカメラにより送られる映像を実時間で表示する。画像は、Choregraphe がカメラと正しく構成された映像キャプチャーモジュールとを有するロボットに接続された時だけ、利用可能である。映像ディスプレイが活性化されると、ユーザは学習をトリガすることができる。次に、カウントダウンが画像上に現れる。このときユーザはカメラの前の対象物を提示するのに例えば4秒を有する。カウントダウンの終わりに、画像が捕捉され記録される。次に、ユーザは、固定画像上にポリゴンを描くことにより画像内の興味のある対象物の輪郭を描かなければならない。ポリゴンが閉じられると、ダイアログが開き、対象物を定義するキーワードを入力するようユーザに依頼する。

40

【0184】

各学習は、Choregraphe によりユーザコンピュータ上に保存されるデータベース内にエントリを生成する。学習が終了すると、ボタンはロボットにデータベースの縮小版を送ることができるようにする。このとき、物体認識モジュールはこのデータベースを使用することになる。対象物が認識されると、関連キーワードを含む事象がロボット上でトリガされる。

【0185】

Choregraphe はまた、ロボットの行動のエディタである。図4に対する注釈

50

として先に説明したように、行動は、ロボットにより実行され得るコンピュータプログラムと同様なオブジェクトである。これらの行動をインストールしロボット上で実行するために、ロボット上の行動を管理するためのインターフェースの開発が行われた。Choregrapheがロボットに接続されると、アプリケーションのメニューのエントリにより行動マネージャを表示できるようにする。これは、これらを実行するための一組のボタンだけでなくロボット上にインストールされる行動のリストも表示するモーダルウィンドウである。

【0186】

インストールされた行動毎に、その名前と、その状態（現在実行中か実行中でないか）と、ロボットが起動されると行動が実行されなければならないかどうかを定義する属性と、が表示される。行動を開始または停止するためには、その現在状態を表示するアイコンをクリックすることで十分であり、その効果は状態をトグル切り替えすることである。行動が終了されると、状態は自動的に元に切り替えられ「停止される」。属性「スタートアップ開始」はチェックマークが付けられるボックスである。これは属性の現在値を示し、ユーザはこの値を変えるためにこれを単純にクリックすることができる。

10

【0187】

行動のリストと共に表示されるボタンは、行動のいくつかを追加すること、行動のいくつかを削除すること、行動のいくつかをユーザコンピュータに転送することができるようにする。したがってユーザは、ロボット上にインストールされた行動を、あたかもそれらが彼のコンピュータ上のファイルかのように、極めて容易に操作することができる。特に、ユーザは、行動を、ロボット上に記録する必要無く、ダウンロードし、修正し、彼のコンピュータ上に再インストールすることができる。

20

【0188】

このときユーザによりインストールされた行動は、時間的一致の制約条件下で、および様々な行動ボックス、行動フレーム、およびタイムラインにより定義された行動間で並行して実行することができる。

【0189】

上に述べた例は本発明の実施形態の例示として与えられた。これは決して本発明の分野を限定するものではなく、本発明の分野は以下の特許請求範囲により規定される。

【 図 4 】

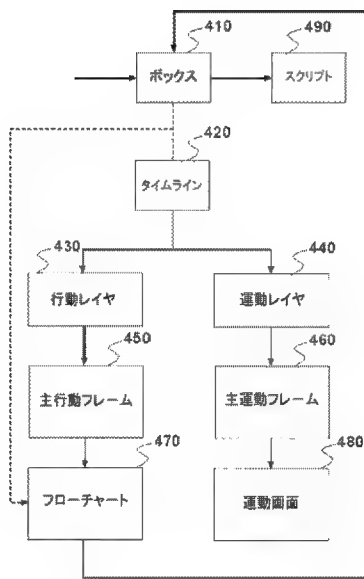


図 4

【 図 5 】

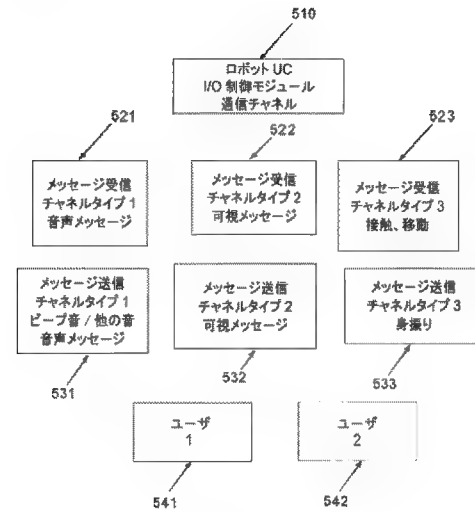


図 5

【 図 6 】

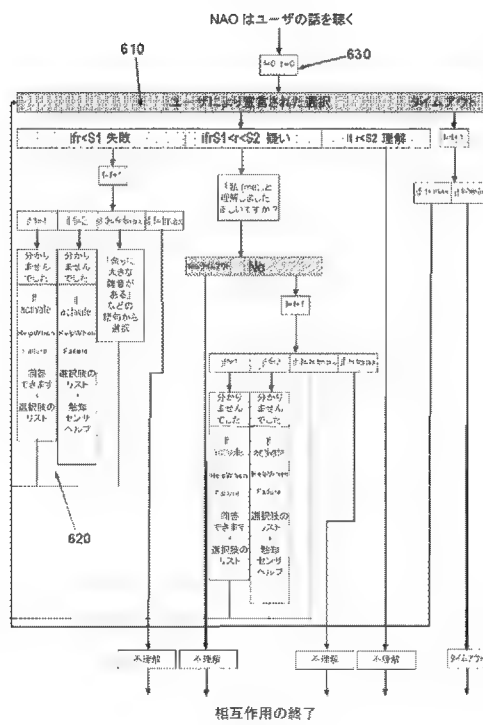


図 6

【 図 7 a 】

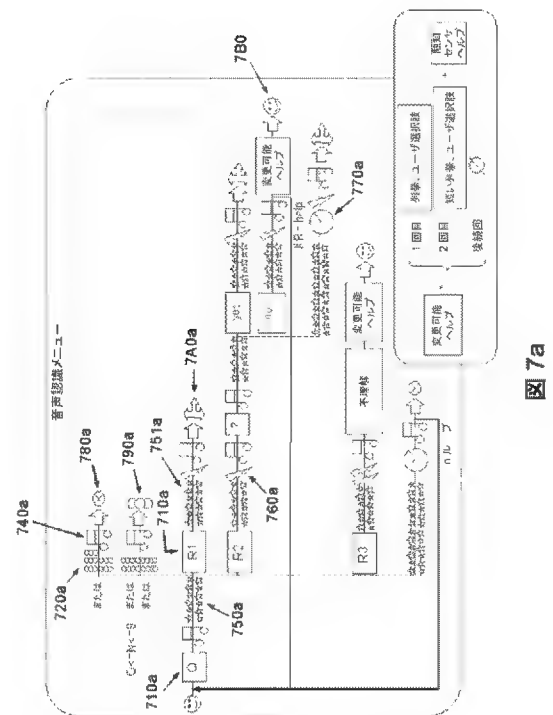


図 7a

【 図 7 b 】

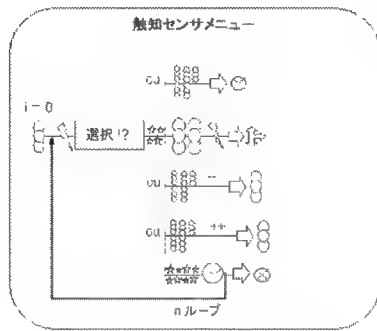


図 7b

【 図 7 c 】

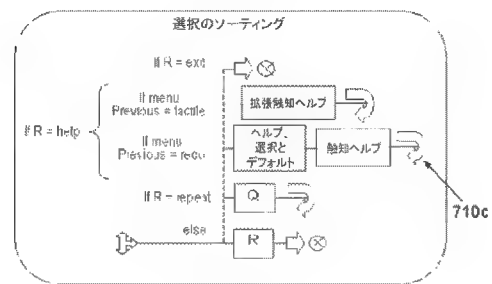


図 7c

【 図 8 a 】

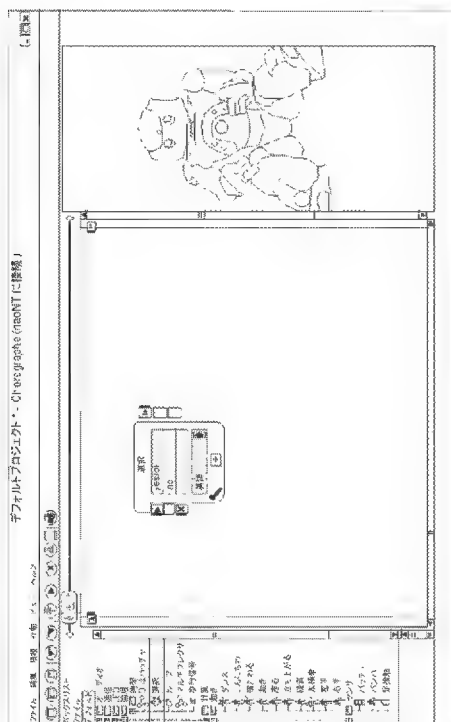


図 8a

【 図 8 b 】

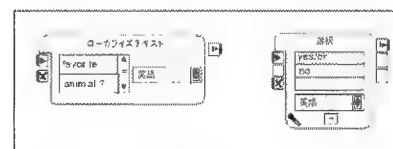


図 8b

【 図 8 c 】

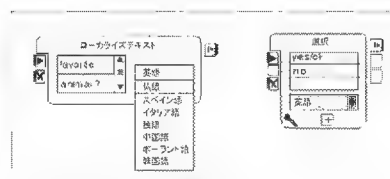


図 8c

【 図 8 d 】

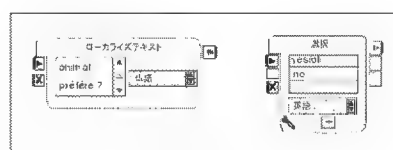


図 8d

【 図 8 e 】

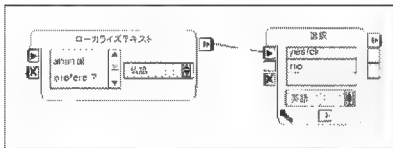


図 8e

【 図 9 a 】

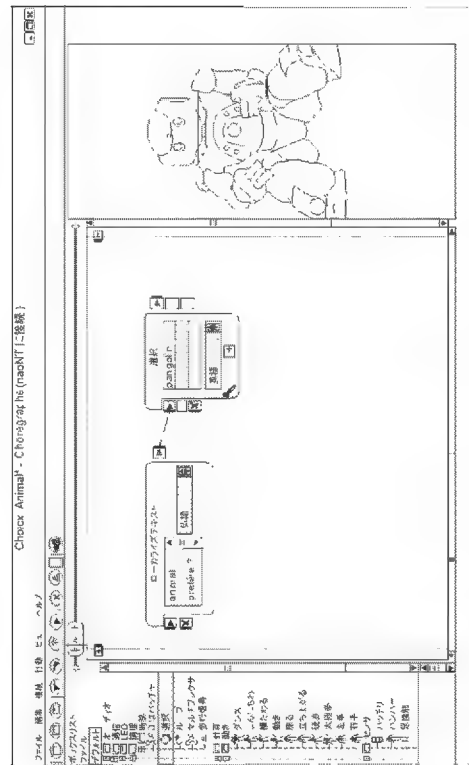


図 9a

【 図 9 b 】

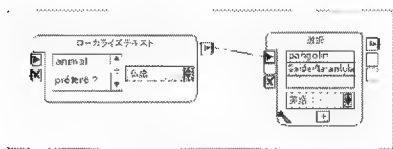


図 9b

【 図 9 e 】

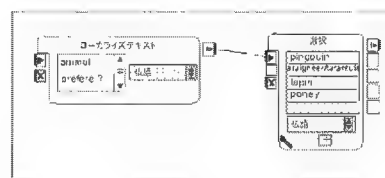


図 9e

【 図 9 c 】

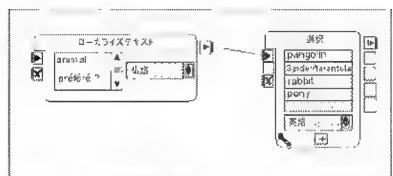


図 9c

【 図 9 d 】

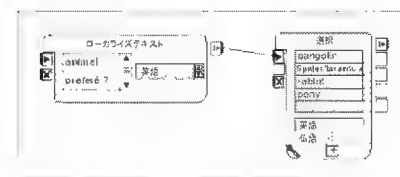


図 9d

【 図 1 0 a 】

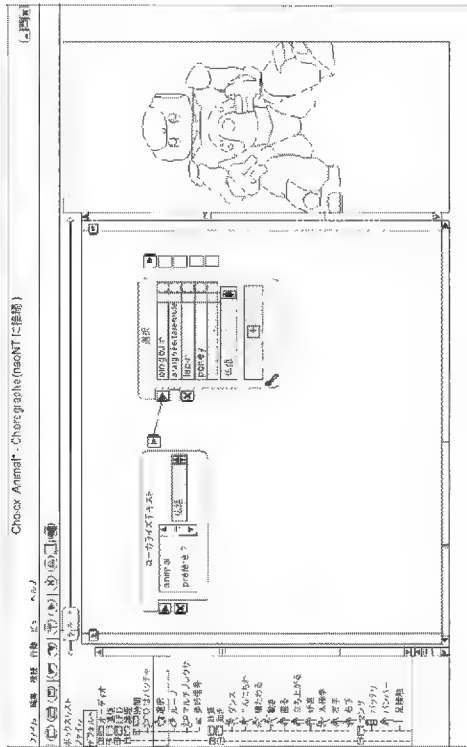


図 10a

【 図 1 0 b 】

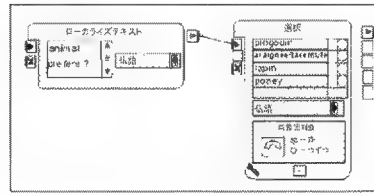


図 10b

【 図 1 0 c 】

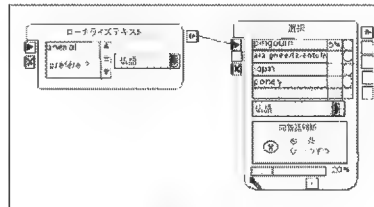


図 10c

【 図 1 0 d 】

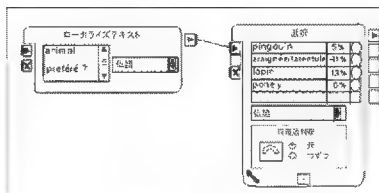


図 10d

【 図 1 1 a 】

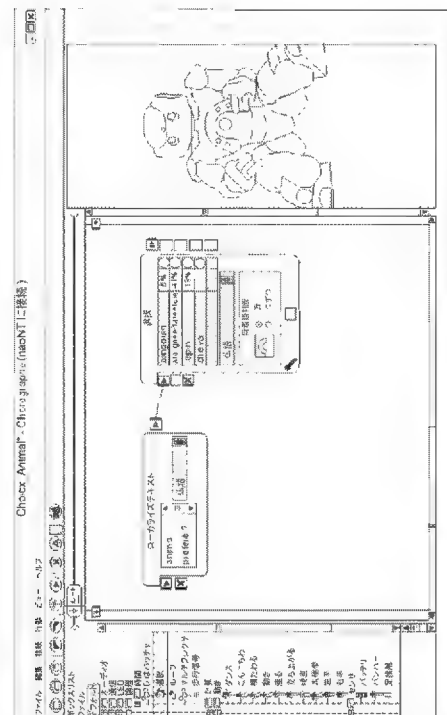


図 11a

【 図 1 1 b 】

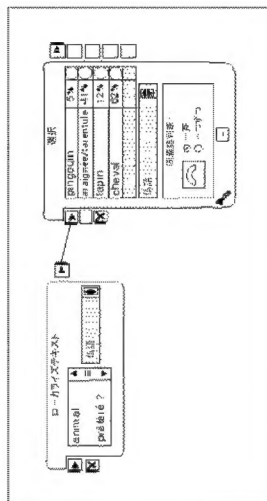


図 11b

【 図 1 2 a 】

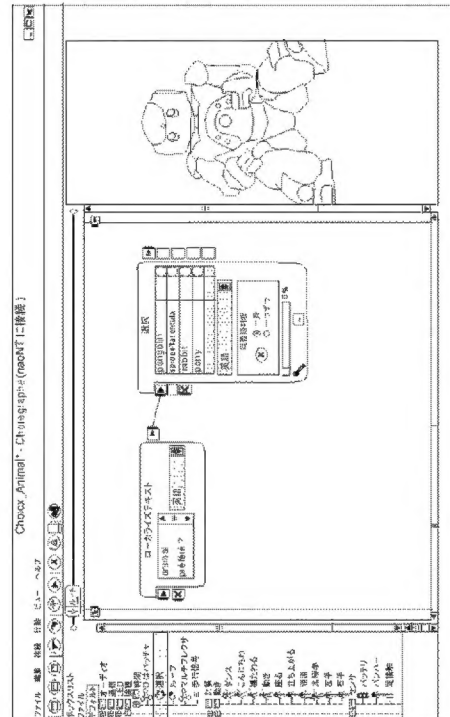


図 12a

【 図 1 2 b 】

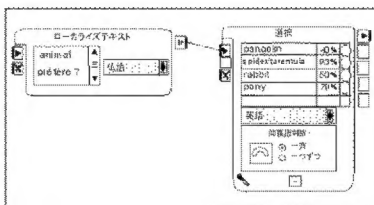


図 12b

【 図 1 2 d 】

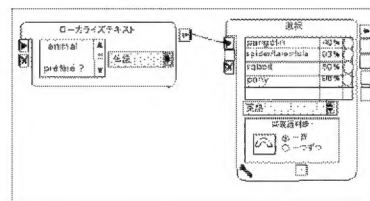


図 12d

【 図 1 2 c 】

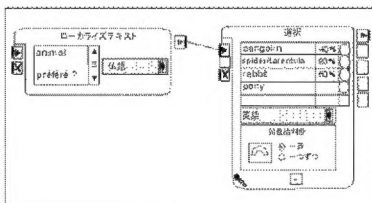


図 12c

【 図 1 3 a 】

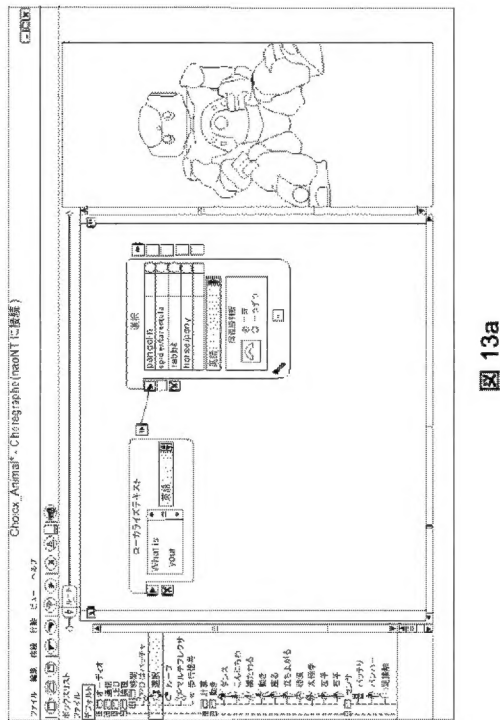


図 13a

【 図 1 3 b 】

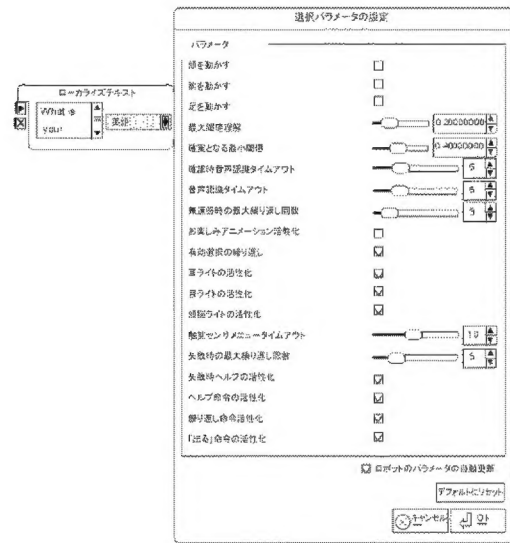


図 13b

【 図 1 3 c 】

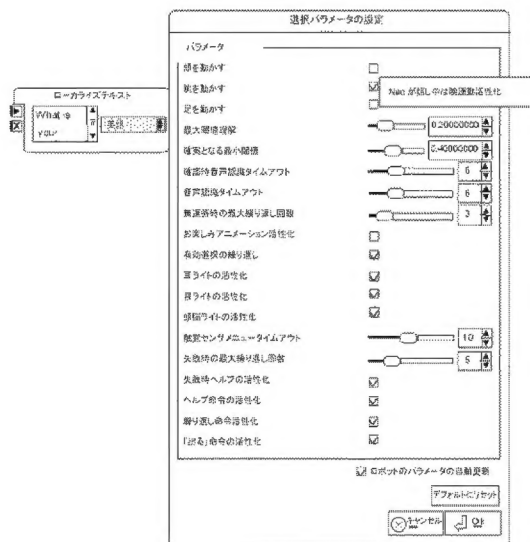


図 13c

【 図 1 3 d 】

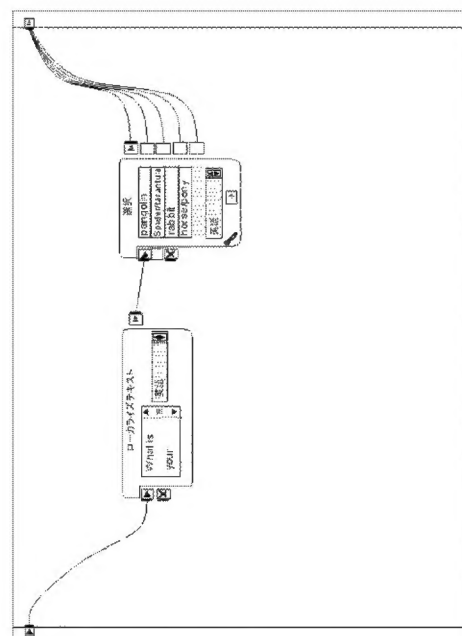


図 13d

【 図 1 】

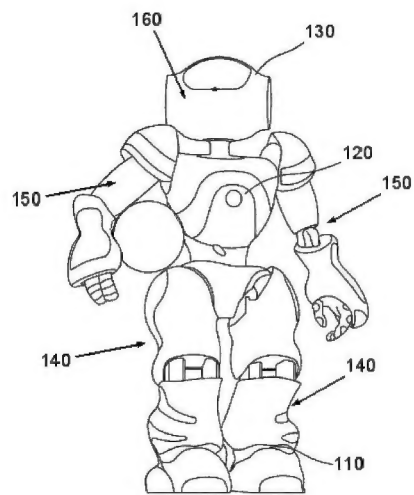


FIG.1

【 図 2 a 】

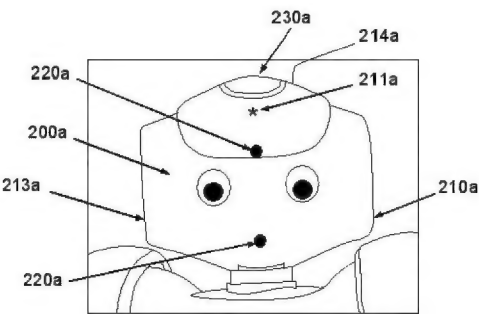


FIG.2a

【 図 2 b 】

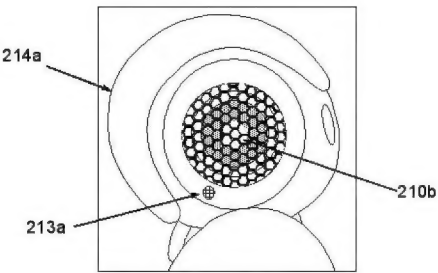


FIG.2b

【 図 3 】

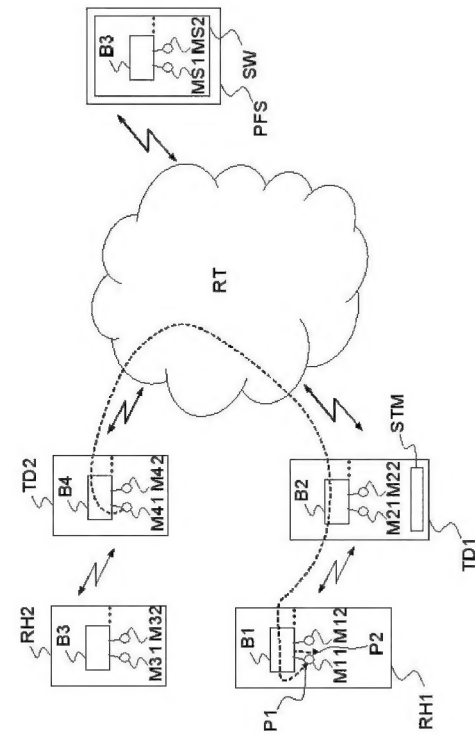


FIG.3

フロント ページの続き

F ターム(参考) 3C707 BS27 CS08 JS03 MT14 WA03 WA13 WL07 WM07 WM25 WM30
5E555 AA26 AA46 AA47 AA61 BA01 BA38 BA90 BB01 BB38 BB40
BC04 BC10 BE20 CA11 CA18 CA31 CA41 CA42 CA45 CA47
CB02 CB64 CB66 CC01 DA01 DA22 DA23 DA24 DA30 DB03
DB32 DB60 DC84 DD01 DD06 EA14 EA22 EA23 FA00